



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Coexistence of MIMO Radar and FD MIMO Cellular Systems with QoS Considerations

Citation for published version:

Biswas, S, Singh, K, Taghizadeh, O & Ratnarajah, T 2018, 'Coexistence of MIMO Radar and FD MIMO Cellular Systems with QoS Considerations', *IEEE Transactions on Wireless Communications*, pp. 1-1.
<https://doi.org/10.1109/TWC.2018.2866044>

Digital Object Identifier (DOI):

[10.1109/TWC.2018.2866044](https://doi.org/10.1109/TWC.2018.2866044)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

IEEE Transactions on Wireless Communications

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Coexistence of MIMO Radar and FD MIMO Cellular Systems with QoS Considerations

Sudip Biswas, Keshav Singh, Omid Taghizadeh, and Tharmalingam Ratnarajah

Abstract—In this work, the feasibility of spectrum sharing between a multiple-input multiple-output (MIMO) radar system (RS) and a MIMO cellular system (CS), comprising of a full duplex (FD) base station (BS) serving multiple downlink and uplink users at the same time and frequency is investigated. While a joint transceiver design technique at the CS's BS and users is proposed to maximise the probability of detection (PoD) of the MIMO RS, subject to constraints of quality of service (QoS) of users and transmit power at the CS, null-space based waveform projection is used to mitigate the interference from RS towards CS. In particular, the proposed technique optimises the performance of PoD of RS by maximising its lower bound, which is obtained by exploiting the monotonically increasing relationship of PoD and its non-centrality parameter. Numerical results show the utility of the proposed spectrum sharing framework, but with certain trade-offs in performance corresponding to RS's transmit power, RS's PoD, CS's residual self interference power at the FD BS and QoS of users.

Index Terms—Multiple-input multiple-output (MIMO), full-duplex (FD), spectrum sharing, MIMO radar, quality-of-service (QoS), transceiver design, convex optimization.

I. INTRODUCTION

The proliferation of wireless devices and services along with static spectrum allocation have resulted in the dearth of spectral resources, which has empowered the vision of spectrum sharing between federal incumbents, such as radar (maritime, surveillance, weather, etc.) and commercial communication systems, such as cellular systems (CSs) [1], [2]. As a result, coexistence between radar and communication systems through technologies, such as Cognitive Radios (CRs) [3], Licensed/Authorized Shared Access (LSA/ASA) [4], [5] and Spectrum Access Systems (SAS) [6] has captured the attention of both academia and industry. However, irrespective of the technology being used, spectrum sharing between radar and communication systems brings a new set of challenges into picture, such as the *harmful interferences generated by CSs towards the radar* and vice-versa, which can potentially degrade the quality-of-service (QoS) of both systems.

In light of this, while the authors in [7] addressed the possibility of spectrum sharing between a rotating radar and

CS, the problem of spectrum sharing between a pulsed, search radar (primary) and 802.11 WLAN (secondary) was investigated in [8]. Similarly, an opportunistic spectrum access approach was proposed in [9] for spectrum sharing between rotating radar and CS. In this approach, the communication system is permitted to transmit signals *if and only if* the space and frequency spectra are not being utilised by the radar, hence prohibiting simultaneous operation of the radar and communication system. To alleviate this issue, LSA/ASA was studied in [4], [5], which allows for the channel state information (CSI) of both systems to be accessible to each other to an extent, where by techniques, such as null-space projection (NSP) of waveforms [10], beamforming [11], etc., can be utilised to mitigate interference at the coexisted systems. Furthermore, to increase waveform diversity and attain higher detection probability, the rotating radar was extended to multiple-input multiple-output (MIMO) radar in [14]–[17]. Accordingly, while a MIMO radar operating in the presence of clutter and a MIMO CS was considered in [12], a null-space projection (NSP) technique for waveform design was proposed in [14] to facilitate the coexistence between a MIMO radar and a MIMO CS. Further, a signal processing scheme for a MIMO radar was proposed in [17] for effectively minimising the arbitrary interferences generated by CSs from any direction towards the radar.

Besides spectrum sharing, full-duplex (FD) is another emerging technology that can potentially double the spectrum efficiency of currently deployed half duplex (HD) systems by fully utilising the spectrum resources in the same time and frequency [18], [19]. However, the self-interference (SI) generated by signal leakage from the transmitting antennas to its receiving antennas, dominates the performance of FD systems. However, recent advancements in interference cancellation techniques [20] suggest that SI can be mitigated in the analog and digital domain to the extent that only a residual SI (RSI) is left behind. This RSI is caused by the non-ideal nature of the transmit and receive chains [21], also known as hardware impairments (amplifier non-linearity, phase noise, I/Q channel imbalance, etc.) and can be mitigated through digital beamforming techniques. Another predicament of FD systems is the co-channel interference (CCI) to the downlink (DL) users caused by the transmission from the uplink (UL) users, which has also been shown in literature to be suppressed through digital beamforming and user-clustering techniques. Accordingly, to harvest the full potential of FD, in recent studies [19], [22], [23] RSI and CCI were integrated in the design of transceivers for FD systems. However, most prior works consider perfect channel state information (CSI) at the

Manuscript received January 03, 2018; revised March 24, 2018; accepted August 03, 2018. The associate editor coordinating the review of this paper and approving it for publication was Prof. Shi Jin.

This work was supported by the UK-India Education and Research Initiative Thematic Partnerships under grant number DST-UKIERI-2016-17-0060.

S. Biswas, K. Singh, and T. Ratnarajah are with the Institute for Digital Communications, School of Engineering, University of Edinburgh, Edinburgh, UK (e-mail: {k.singh, sudip.biswas, t.ratnarajah}@ed.ac.uk).

O. Taghizadeh is with the Institute for Theoretical Information Technology, RWTH Aachen University, Germany (e-mail: taghizadeh@ti.rwth-aachen.de).

The corresponding author of this paper is Keshav Singh.

transmitter, which is improbable to attain due to large training overhead and low signal-to-noise-ratio obtained after beam-forming. Hence, it is important to design robust transceivers, which can provide respectable performance for the co-existed systems even with imperfect or limited CSI estimates.

Motivated by the above discussion, in this paper we consider a two-tier spectrum sharing framework, where a multiple-input-multiple-output (MIMO) radar system (RS) is the spectrum sharing entity and a FD MIMO CS is the beneficiary. Unlike existing literature [7]–[17], in this work we consider a framework wherein, by utilising the spectrum of the MIMO RS, the FD CS's BS serves multiple downlink (DL) and uplink (UL) users simultaneously in same time and frequency resources. At this point, we would like to note that though the consideration of FD mode at the CS provides better QoS for the users, it also induces *higher intensity of interference towards the RS, which might affect its probability of detection (PoD)*. Hence, the requirements for the coexistence of a FD CS, with a MIMO RS requires revisiting. Further, the inclusions of imperfect CSI, hardware impairments and CCI in the design of the proposed co-existed network result in a rather arduous problem, which requires rigorous optimisation and analysis and is addressed in this paper. Accordingly, to enable efficient spectrum sharing between a FD MU MIMO CS and MIMO RS, we focus on the design of:

- Beamforming matrices at the FD CS to mitigate interference towards the RS, while also providing data to its users. In particular, we formulate an optimisation problem for maximising the PoD of the RS for a given worst-case channel set, subject to the constraints of QoS at the UL and DL users and powers at the UL users and the FD BS.
- Projection matrix at the MIMO RS to mitigate interference towards the FD CS. In particular, all the interference channels from the RS to CS are combined into one and a null-space is created. The RS's signal is then projected onto the null-space of the combined interference channel.

The following notations are used throughout this paper. Boldface capital and small letters denote matrices and vectors, respectively. While the transpose and the conjugate transpose are denoted by $(\cdot)^T$ and $(\cdot)^H$, respectively, $\|\mathbf{A}\|_F$ and $\|\mathbf{a}\|_2$ denote the Frobenius norm of a matrix \mathbf{A} and the Euclidean norm of a vector \mathbf{a} , respectively. \otimes denotes Kronecker product and \perp denotes the statistical independence. The matrices \mathbf{I}_N and $\mathbf{0}_{M \times N}$ denote a $N \times N$ identity matrix and a $M \times N$ zero matrix, respectively. The notations $\mathbb{E}(\cdot)$ and $\text{tr}(\cdot)$ refer to expectation and trace, respectively and $\text{diag}(\mathbf{A})$ and $\text{vec}(\mathbf{A})$ generate a diagonal matrix with the same diagonal element as \mathbf{A} and a $mn \times 1$ column vector obtained by stacking the columns of the matrix \mathbf{A} on top of one another, respectively.

II. SYSTEM MODEL

We consider the coexistence of a FD multi-user MIMO CS with a MIMO radar system (RS) as shown in Fig. 1, where the CS operates in the spectrum shared by the RS over a bandwidth of B Hz. The CS comprises of a FD MIMO BS, which consists of M_0 transmit and N_0 receive

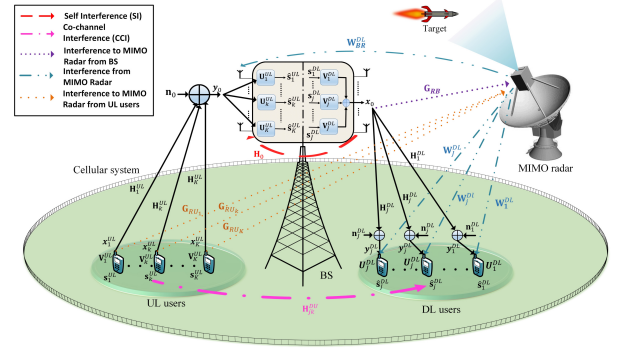


Fig. 1: Spectrum coexistence between FD MIMO CS and MIMO RS.

antennas, and J DL, and K UL users. All DL and UL users operate in half-duplex (HD) mode and each DL and UL user is equipped with N_j receive and M_k transmit antennas, respectively. Furthermore, the MIMO RS is equipped with R_R receive and R_T transmit antennas. In the following subsections, we provide details on the architecture of the two systems.

1) *FD MIMO CS*: As shown in Fig. 1, $\mathbf{H}_k^{UL} \in \mathbb{C}^{N_0 \times M_k}$ and $\mathbf{H}_j^{DL} \in \mathbb{C}^{N_j \times M_0}$ represent the k -th UL and the j -th DL channel, respectively. The SI channel at the FD BS and the CCI channel between the k -th UL and j -th DL users are denoted as $\mathbf{H}_0 \in \mathbb{C}^{N_0 \times M_0}$ and $\mathbf{H}_{jk}^{DU} \in \mathbb{C}^{N_j \times M_k}$, respectively. Finally, the *interference channels* from the FD BS and k th UL user to the RS are denoted as $\mathbf{G}_{RB} \in \mathbb{C}^{R_R \times M_0}$ and $\mathbf{G}_{RU_k} \in \mathbb{C}^{R_R \times M_k}$, respectively. Now, let $\mathbf{s}_k^{UL}(l) \in \mathbb{C}^{d_k^{UL} \times 1}$ and $\mathbf{s}_j^{DL}(l) \in \mathbb{C}^{d_j^{DL} \times 1}$ denote the transmitted symbols by the k -th UL user and the FD BS, respectively, at time index l , with $l = 1, \dots, L$, such that $\mathbb{E}[\mathbf{s}_k^{UL}(l)(\mathbf{s}_k^{UL}(l))^H] = \mathbf{I}_{d_k^{UL}}$ and $\mathbb{E}[\mathbf{s}_j^{DL}(l)(\mathbf{s}_j^{DL}(l))^H] = \mathbf{I}_{d_j^{DL}}$. Here, d_k^{UL} and d_j^{DL} denote the data streams from the k th UL user and for the j th DL user, respectively, and L is the total number of time samples used for CS's communication. The symbols \mathbf{s}_k^{UL} and \mathbf{s}_j^{DL} are first precoded by matrices $\mathbf{V}_k^{UL} \in \mathbb{C}^{M_k \times d_k^{UL}}$ and $\mathbf{V}_j^{DL} \in \mathbb{C}^{M_0 \times d_j^{DL}}$, respectively, such that the signals transmitted from the k -th UL user and the FD BS at time index l are given as $\mathbf{x}_k^{UL}(l) = \mathbf{V}_k^{UL} \mathbf{s}_k^{UL}(l)$ and $\mathbf{x}_0(l) = \sum_{j=1}^J \mathbf{V}_j^{DL} \mathbf{s}_j^{DL}(l)$, respectively. Further, similar to [24], we model the imperfections in the transmitter/receiver chains (oscillators, analog-to-digital converters (ADCs), digital-to-analog converters (DACs), and power amplifiers) as an additive Gaussian noise. In particular, we consider an additive white Gaussian term as “transmitter noise” (“receiver distortion”) at each transmit (receive) antenna, whose variance is $\psi(v)$ times the power of the undistorted signal at the corresponding chain.

2) *MIMO RS*: Let \mathbf{W}_i denote the *interference channels* shared by the MIMO RS with the CS, where $i = 1, \dots, \mathcal{I}$, with $\mathcal{I} = 1 + J$. As shown in Fig. 1, $\mathbf{W}_{BR}^{DL} \in \mathbb{C}^{N_0 \times R_T}$ and $\mathbf{W}_{jR}^{DL} \in \mathbb{C}^{N_j \times R_T}$ denotes the interference channels from RS's transmitter to CS's BS and j -th DL user, respectively, with $\{\mathbf{W}_{BR}^{DL}, \mathbf{W}_{jR}^{DL}\} \subseteq \mathbf{W}, \forall j$. Let $\mathbf{s}_R(l) \in \mathbb{C}^{R_T \times 1}$ represents the transmitted symbol by the RS at time index l , with $l = 1, \dots, L_R$, and $\mathbb{E}[\mathbf{s}_R(l)(\mathbf{s}_R(l))^H] = \mathbf{I}_{R_T}$. Here,

L_R is the total number of time samples used for the RS's communication. For the ease of derivation, hereinafter we assume that the time duration of the RS's waveform is the same as the communication signals with $L_R = L$ [11].

3) *CSI acquisition*: Acquiring CSI at both systems is important to ensure an interference free communication. In this work, we assume that some amount of CSI, if not full is available at the communication nodes¹. For the CS, providing its CSI to the RS is incentivised by the promise of zero interference from the RS. On the other hand, it is more challenging to obtain an accurate estimate of the CSI of the RS at the CS, as the RS might not be willing to cooperate with the CS owing to security concerns. Hence, it might not be possible to obtain a full CSI at the CS and only partial CSI may be obtained through techniques such as blind environmental learning [25], realisation of a band manager with the authority of exchanging CSI between the RS and CS [26], etc. Hence, to model the imperfections caused due to imperfect CSI, in this paper we will consider a norm-bounded channel estimation error model for the links between the CS and RS.

III. SPECTRUM SHARING MIMO RS

As discussed in Section I, the considered sharing framework requires both the RS and CS to work in tandem to mitigate interference towards each other. Accordingly, in this section, we formulate the part played by the MIMO RS for efficient spectrum sharing between the two systems. The goal of the RS is to map $\mathbf{s}_R(l)$ onto the null-space of \mathbf{W} in order to avoid interference towards the CS. Accordingly, the symbols \mathbf{s}_R are first multiplied by a projector matrix $\mathbf{P} \in \mathbb{C}^{R_T \times R_T}$, such that the signal transmitted by the RS is given by $\mathbf{x}_R(l) = \mathbf{P}\mathbf{s}_R(l)$.

1) *Target detection*: In order to detect a target in the far field, we apply a binary hypothesis and choose between two cases: 1) \mathcal{H}_0 : No target but active CS and 2) \mathcal{H}_1 : Both the target and the CS are active. Further, in this work we consider only a single target with no interference from other sources, but the CS in order to study the impact of CS's interference on PoD of RS.

Accordingly, the hypothesis testing problem for an echo wave in a single range-Doppler bin of the RS can be written as² in (1) as shown on the top of the next page, where \mathcal{H}_1 comprises of the discrete time signal vector received by RS at an angle ϕ , and \mathcal{H}_0 comprises of the interference plus noise signals from the CS only. In the above, P_R denotes the transmit power of RS, α_r indicates the complex path loss exponent of the radar-target-radar path including the propagation loss and the coefficient of reflection, $\mathbf{n}_R(l) \sim \mathcal{CN}(0, \sigma_R^2 \mathbf{I}_R)$, and $\mathbf{A}(\phi)$ denotes the transmit-receive steering matrix defined as

$$\mathbf{A}(\phi) \triangleq \mathbf{a}_R(\phi) \mathbf{a}_T^T(\phi). \quad (2)$$

Here, $\mathbf{a}_T \in \mathbb{C}^{R_T \times 1}$ and $\mathbf{a}_R \in \mathbb{C}^{R_R \times 1}$ denote the transmit and receive steering vectors of RS's antenna array. Hereinafter, we assume that $R_R = R_T = R$ and $\mathbf{a}_R(\phi) = \mathbf{a}_T(\phi) = \mathbf{a}(\phi)$.

Accordingly, the i -th element at the r -th column of the matrix $\mathbf{A}(\phi)$, can be written as

$$A_{ir}(\phi) = \exp\left(-j \frac{2\pi}{\lambda} [\sin(\phi); \cos(\phi)]^T (\mathbf{z}_i + \mathbf{z}_r)\right), \quad (3)$$

where $\mathbf{z}_i = [z_i^1; z_i^2]$ is the location of the i -th element of the antenna array and λ is the wavelength of the carrier. Since, the deterministic parameters α_r and ϕ are unknown³, we adopt the generalised likelihood ratio test (GLRT) [27], which has the advantage of replacing the unknown parameters with their maximum likelihood (ML) estimates for determining the probability of detection (PoD). Hereinafter, unless otherwise stated, we drop the time index l in this section for notational convenience. The sufficient statistic of the received signal can then be found using matched filtering as [28]

$$\begin{aligned} \hat{\mathbf{Y}} &= \frac{1}{\sqrt{L}} \sum_{l=1}^L \mathbf{y}_R \mathbf{x}_R^H = \alpha_r \sqrt{L P_R} \mathbf{A}(\phi) \mathbf{P} + \frac{1}{\sqrt{L}} \\ &\times \sum_{l=1}^L \left(\mathbf{G}_{RB} \left(\sum_{j=1}^J \mathbf{V}_j^{DL} \mathbf{s}_j^{DL} + \mathbf{c}_0 \right) \right. \\ &\quad \left. + \sum_{k=1}^K \mathbf{G}_{RU_k} (\mathbf{V}_k^{UL} \mathbf{s}_k^{UL} + \mathbf{c}_k^{UL}) + \mathbf{n}_R \right) \mathbf{x}_R^H. \end{aligned} \quad (4)$$

From (4), the vectorization of $\hat{\mathbf{Y}}$ can be written as

$$\begin{aligned} \hat{\mathbf{y}} &= \text{vec}(\hat{\mathbf{Y}}) \\ &= \alpha_r \sqrt{L P_R} \text{vec}(\mathbf{A}(\phi) \mathbf{P}) \\ &\quad + \text{vec}\left(\frac{1}{\sqrt{L}} \sum_{l=1}^L \left(\mathbf{G}_{RB} \left(\sum_{j=1}^J \mathbf{V}_j^{DL} \mathbf{s}_j^{DL} + \mathbf{c}_0 \right) \right. \right. \\ &\quad \left. \left. + \sum_{k=1}^K \mathbf{G}_{RU_k} (\mathbf{V}_k^{UL} \mathbf{s}_k^{UL} + \mathbf{c}_k^{UL}) + \mathbf{n}_R \right) \mathbf{x}_R^H \right) \\ &\triangleq \alpha_r \sqrt{L P_R} \text{vec}(\mathbf{A}(\phi) \mathbf{P}) + \boldsymbol{\Psi}, \end{aligned} \quad (5)$$

where $\boldsymbol{\Psi}$ is zero-mean, complex Gaussian distributed, and has a non-white covariance matrix of

$$\boldsymbol{\chi} = \begin{pmatrix} \mathbf{P}^H (\tilde{\boldsymbol{\chi}} + \sigma_R^2 \mathbf{I}_R) \mathbf{P} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{P}^H (\tilde{\boldsymbol{\chi}} + \sigma_R^2 \mathbf{I}_R) \mathbf{P} \end{pmatrix}. \quad (6)$$

In the above, $\boldsymbol{\chi} \in \mathbb{C}^{R^2 \times R^2}$ and $\tilde{\boldsymbol{\chi}} = \left(\mathbf{G}_{RB} \sum_{j=1}^J \left(\mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H + \psi \text{diag}(\mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H) \right) \right) (\mathbf{G}_{RB})^H + \sum_{k=1}^K \left(\mathbf{G}_{RU_k} \mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H + \psi \text{diag}(\mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H) \right) (\mathbf{G}_{RU_k})^H$. However, the GLRT in [28] was applied in the presence of white noise only. Hence, to convert the covariance matrix in (6) into white, we apply a whitening filter $\boldsymbol{\Pi}^H$, obtained after the Cholesky decomposition⁴ of $\boldsymbol{\chi}^{-1}$ as $\boldsymbol{\chi}^{-1} = \boldsymbol{\Pi} \boldsymbol{\Pi}^H$, with $\boldsymbol{\Pi}$ being a lower triangle matrix. Now, the hypothesis testing problem in (1) can be equivalently rewritten as

$$\hat{\mathbf{y}} = \begin{cases} \mathcal{H}_1 : \alpha_r \sqrt{L P_R} \boldsymbol{\Pi}^H \mathbf{A}(\phi) + \boldsymbol{\Pi}^H \boldsymbol{\Psi}, \\ \mathcal{H}_0 : \boldsymbol{\Pi}^H \boldsymbol{\Psi}, \end{cases} \quad (7)$$

¹CSI estimation can be performed via the exchange of training sequences and feedback, and the application of usual CSI estimation methods [19].

²For simplicity, we ignore the interference from clutter and false targets.

³Due to the availability of CS's CSI at the RS, we assume that the covariance matrix of the interference-plus-noise has been accurately estimated by the RS.

⁴Note that $\boldsymbol{\chi}$ and $\boldsymbol{\chi}^{-1}$ are both positive-definite Hermitian matrices.

$$\mathbf{y}_R(l) = \begin{cases} \mathcal{H}_1 : \alpha\sqrt{P_R}\mathbf{A}(\phi)\mathbf{P}\mathbf{s}_R(l) + \mathbf{G}_{RB} \left(\overbrace{\sum_{j=1}^J \mathbf{V}_j^{DL}\mathbf{s}_j^{DL}(l) + \mathbf{c}_0}^{\text{Interference from BS}} \right) \\ \quad + \underbrace{\sum_{k=1}^K \mathbf{G}_{RU_k} (\mathbf{V}_k^{UL}\mathbf{s}_k^{UL}(l) + \mathbf{c}_k^{UL})}_{\text{Interference from UL users}} + \mathbf{n}_R(l), & 1 \leq l \leq L, \\ \mathcal{H}_0 : \mathbf{G}_{RB} \left(\sum_{j=1}^J \mathbf{V}_j^{DL}\mathbf{s}_j^{DL}(l) + \mathbf{c}_0 \right) + \sum_{k=1}^K \mathbf{G}_{RU_k} (\mathbf{V}_k^{UL}\mathbf{s}_k^{UL}(l) + \mathbf{c}_k^{UL}) + \mathbf{n}_R(l), & 1 \leq l \leq L, \end{cases} \quad (1)$$

where $\bar{\mathbf{A}}(\phi) = \text{vec}(\mathbf{A}(\phi)\mathbf{P})$. If $p(\hat{\mathbf{y}}, \hat{\alpha}_r, \hat{\phi}, \mathcal{H}_1)$ and $p(\hat{\mathbf{y}}, \mathcal{H}_0)$ denote the probability density function under \mathcal{H}_1 and \mathcal{H}_0 , respectively, and $\hat{\alpha}_r$ and $\hat{\phi}$ indicate the ML estimation of α_r and ϕ under hypothesis \mathcal{H}_1 , which is expressed as $[\hat{\alpha}_r, \hat{\phi}] = \max_{\alpha_r, \phi} p(\hat{\mathbf{y}} | \alpha_r, \phi, \mathcal{H}_1)$, then the GLRT can be given by

$$\begin{aligned} \ln L_{\hat{\mathbf{y}}}(\hat{\phi}_{ML}) &= \frac{|\bar{\mathbf{A}}^H(\hat{\phi}_{ML}) \mathbf{\Pi} \mathbf{\Pi}^H \hat{\mathbf{y}}|^2}{\|\mathbf{\Pi}^H \bar{\mathbf{A}}^H(\hat{\phi}_{ML})\|^2} \leq \bar{\delta}, \\ &= \frac{|\text{tr}(\hat{\mathbf{Y}} \mathbf{P}^H \mathbf{A}^H(\hat{\phi}_{ML}) \hat{\chi}^{-1})|^2}{\text{tr}(\mathbf{A}(\hat{\phi}_{ML}) \mathbf{P} \mathbf{P}^H \mathbf{A}^H(\hat{\phi}_{ML}) \hat{\chi}^{-1})} \underset{\mathcal{H}_0}{\leq} \underset{\mathcal{H}_1}{\bar{\delta}}, \end{aligned} \quad (8)$$

where $\bar{\delta}$ denotes the decision threshold and $\hat{\chi} = \tilde{\chi} + \sigma_R^2 \mathbf{I}_R$. According to [29], the asymptotic statistic of $L_{\hat{\mathbf{y}}}(\hat{\phi}_{ML})$ for both the hypothesis is written as

$$\ln L_{\hat{\mathbf{y}}}(\hat{\phi}_{ML}) \sim \begin{cases} \mathcal{H}_1 : \chi^2_2(\rho), \\ \mathcal{H}_0 : \chi^2_2, \end{cases} \quad (9)$$

where χ^2_2 denotes the central chi-squared distributions with two degrees of freedom (DoFs), $\chi^2_2(\rho)$ is the non-central chi-squared distributions with two DoFs, and ρ indicates the non-central parameter given as

$$\begin{aligned} \rho &= |\alpha_r|^2 LP_R \text{vec}^H(\mathbf{A}(\phi)\mathbf{P}) \chi^{-1} \text{vec}(\mathbf{A}(\phi)\mathbf{P}) \\ &= \Gamma_R \sigma_R^2 \text{tr}(\mathbf{A}(\phi) \mathbf{P} \mathbf{P}^H \mathbf{A}^H(\phi) \hat{\chi}^{-1}). \end{aligned} \quad (10)$$

where $\Gamma_R = |\alpha_r|^2 LP_R / \sigma_R^2$. The decision threshold $\bar{\delta}$ is set according to a desired probability of false alarm P_{FA} as

$$\bar{\delta} = \mathfrak{F}_{\chi^2_2}^{-1}(1 - P_{FA}), \quad (11)$$

where $\mathfrak{F}_{\chi^2_2}^{-1}$ denotes the inverse central chi-squared distribution function with two DoFs. The PoD for the MIMO RS can now be given as

$$P_D = 1 - \mathfrak{F}_{\chi^2_2(\rho)}(\mathfrak{F}_{\chi^2_2}^{-1}(1 - P_{FA})), \quad (12)$$

where $\mathfrak{F}_{\chi^2_2(\rho)}$ is the non-central chi-squared distribution function with two DoFs.

2) *Projection matrix design at MIMO RS*: In order to mitigate the interference towards the CS, we need to design the projection matrix \mathbf{P} in such a way that it projects onto the null space of the interference channels \mathbf{W} . Accordingly, to find the null space of \mathbf{W} , we perform its singular value decomposition (SVD), which can be given as $\mathbf{W} = \mathbf{R}\mathbf{\Omega}\mathbf{X}^H$, where \mathbf{R} and \mathbf{X} are unitary matrices and $\mathbf{\Omega}$ is a diagonal

matrix whose elements are the singular values of \mathbf{W} . Now, let

$$\bar{\mathbf{\Omega}} \triangleq \text{diag}(\bar{\omega}_1 \dots \bar{\omega}_p), \quad (13)$$

where $p \triangleq \min(N_{BS+Users}, R_T)$, $\bar{\omega}_1 > \bar{\omega}_2 > \dots \bar{\omega}_q = \bar{\omega}_{q+1} = \dots = \bar{\omega}_p = 0$ and

$$\tilde{\mathbf{\Omega}} \triangleq \text{diag}(\tilde{\omega}_1 \dots \tilde{\omega}_{R_T}), \quad (14)$$

where $\tilde{\omega}_r \triangleq 0, \forall r \leq q$, $\tilde{\omega}_r \triangleq 1, \forall r > q$, with $\tilde{\mathbf{\Omega}}\tilde{\mathbf{\Omega}} = \mathbf{0}$. Using the above definitions, the beamforming matrix for NSP can be defined as [14]

$$\mathbf{P} \triangleq \mathbf{X}\tilde{\mathbf{\Omega}}\mathbf{X}^H. \quad (15)$$

Proposition 1: When $R_T \gg (N_0 + JN_j)$, the beamforming matrix $\mathbf{P} \in \mathbb{C}^{R_T \times R_T}$ can be projected orthogonally onto the null-space of the entire CS involving the full set of interference channels $\mathbf{W} \in \mathbb{C}^{N_{BS+Users} \times R_T}$.

Proof: The proof is given in Appendix A. ■

IV. SPECTRUM UTILISING FD MIMO CS

After establishing the role of the MIMO RS, in this section, we formulate the part played by the FD MIMO CS in the proposed spectrum sharing framework.

1) *Achievable rate*: Using the spectrum of MIMO RS, the signals received at the FD BS and the j -th DL user of the CS at time index l can be written, respectively as

$$\begin{aligned} \mathbf{y}_0(l) &= \sum_{k=1}^K \mathbf{H}_k^{UL} (\mathbf{x}_k^{UL}(l) + \mathbf{c}_k^{UL}(l)) + \mathbf{H}_0 (\mathbf{x}_0(l) + \mathbf{c}_0(l)) \\ &\quad + \underbrace{\mathbf{e}_0(l) + \sqrt{P_R} \mathbf{W}_{BR}^{DL} \mathbf{s}_R(l)}_{\text{Interference from RS}} + \mathbf{n}_0(l), \end{aligned} \quad (16)$$

$$\begin{aligned} \mathbf{y}_j^{DL}(l) &= \mathbf{H}_j^{DL} (\mathbf{x}_0(l) + \mathbf{c}_0(l)) + \sum_{k=1}^K \mathbf{H}_{jk}^{DU} (\mathbf{x}_k^{UL}(l) + \mathbf{c}_k^{UL}(l)) \\ &\quad + \underbrace{\mathbf{e}_j^{DL}(l) + \sqrt{P_R} \mathbf{W}_j^{DL} \mathbf{s}_R(l)}_{\text{Interference from RS}} + \mathbf{n}_j^{DL}(l), \end{aligned} \quad (17)$$

where, $\mathbf{n}_0(l) \in \mathbb{C}^{N_0}$ and $\mathbf{n}_j^{DL}(l) \in \mathbb{C}^{N_j}$ denote the additive white Gaussian noise (AWGN) vector with zero mean and covariance matrix $\mathbf{R}_0 = \sigma_0^2 \mathbf{I}_{N_0}$ and $\mathbf{R}_j^{DL} = \sigma_j^2 \mathbf{I}_{N_j}$ at the BS and the j -th DL user, respectively, and $\mathbf{c}_k^{UL}(l) \sim \mathcal{CN}(\mathbf{0}, \psi \text{diag}(\mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H))$, $\mathbf{c}_k^{UL}(l) \perp \mathbf{x}_k^{UL}(l)$ and $\mathbf{e}_j^{DL}(l) \sim \mathcal{CN}(\mathbf{0}, v \text{diag}(\mathbf{\Phi}_j^{DL}))$, $\mathbf{e}_j^{DL}(l) \perp \hat{\mathbf{u}}_j^{DL}(l)$ are the transmit and receive distortion at the k -th UL and j -th DL user, respectively with $\psi \ll 1$ and $v \ll 1$. Here, $\mathbf{\Phi}_j^{DL} = \text{Cov}\{\hat{\mathbf{u}}_j^{DL}(l)\}$ and $\hat{\mathbf{u}}_j^{DL}(l) = \mathbf{y}_j^{DL}(l) - \mathbf{e}_j^{DL}(l)$. The transmitter/receiver distortion $\mathbf{c}_0(l)/\mathbf{e}_0(l)$ can be defined

similarly. Like before, we drop the time index l henceforth, unless otherwise stated.

Remark 1: Since the MIMO RS transmits NSP based signals, the interference terms in (16), (17) $\mathbf{W}_{BR}^{DL} \mathbf{s}_R(l)$ and $\mathbf{W}_j^{DL} \mathbf{s}_R(l)$ will effectively be $\mathbf{W}_{BR}^{DL} \mathbf{x}_R(l)$ and $\mathbf{W}_j^{DL} \mathbf{x}_R(l)$ and hence, will be mitigated. However, they have been retained here for tractability of the corresponding expressions, but will be ignored in calculation of the numerical results.

Remark 2: Since the codeword \mathbf{x}_0 and the SI channel \mathbf{H}_0 are known to the BS (its own transmitted signal), the term $\mathbf{H}_0 \mathbf{x}_0$ in (16) can be cancelled out and thus, the remaining part $\mathbf{H}_0 \mathbf{c}_0$ can be treated as the RSI. However, the term $\mathbf{H}_0 \mathbf{x}_0$ has been retained for tractability, but will be ignored in calculation of the numerical results.

Lemma 1: The approximated aggregate interference-plus-noise terms at the k -th UL and j -th DL user, can be given respectively as⁵ in (18) and (19) as shown on the top of the next page.

Proof: This approximation can be easily proofed from (16) and (17) by considering $\psi \ll 1$ and $v \ll 1$ and ignoring the terms ψv . ■

Hence, from (16), (17) and Lemma 1, the achievable rate for the k -th UL user at the BS and the j -th DL user can be given as

$$R_k^{UL} = \log_2 \left| \mathbf{I}_k^{UL} + (\mathbf{U}_k^{UL})^H \mathbf{H}_k^{UL} \mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H (\mathbf{H}_k^{UL})^H \times \mathbf{U}_k^{UL} \left((\mathbf{U}_k^{UL})^H \mathbf{\Sigma}_k^{UL} \mathbf{U}_k^{UL} \right)^{-1} \right|, \quad (20)$$

$$R_j^{DL} = \log_2 \left| \mathbf{I}_j^{DL} + (\mathbf{U}_j^{DL})^H \mathbf{H}_j^{DL} \mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H (\mathbf{H}_j^{DL})^H \times \mathbf{U}_j^{DL} \left((\mathbf{U}_j^{DL})^H \mathbf{\Sigma}_j^{DL} \mathbf{U}_j^{DL} \right)^{-1} \right|. \quad (21)$$

2) *Beamformer design at FD MIMO CS:* The main objective of the CS is to provide the cellular users with data by utilising the spectrum of the RS, but without affecting the PoD of RS. Hence, we formulate the beamforming design problem as

$$\begin{aligned} (\mathbf{P0}) \quad & \max_{\mathbf{V}} \quad P_D \\ \text{subject to } & (C.1) \quad R_k^{UL} \geq R_{k,min}^{UL}, \quad k = 1, \dots, K, \\ & (C.2) \quad R_j^{DL} \geq R_{j,min}^{DL}, \quad j = 1, \dots, J, \\ & (C.3) \quad \text{tr} \left\{ \mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H \right\} \leq P_k, \quad k = 1, \dots, K, \\ & (C.4) \quad \sum_{j=1}^J \text{tr} \left\{ \mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H \right\} \leq P_0, \end{aligned} \quad (22)$$

where $R_{k,min}^{UL}$ in (C.1) and $R_{j,min}^{DL}$ in (C.2) are the minimum QoS requirements for the k -th UL and j -th DL users, respectively. Further, the constraints (C.3) and (C.4) regulate the transmit powers at the k -th UL user and the BS, respectively and $\mathbf{V} = \{\mathbf{V}_k^{UL}, \mathbf{V}_j^{DL}\}$ denotes the set of all transmit beamforming matrices. Since P_D is a monotonically

increasing function with respect to the non-central parameter ρ [29], we can equivalently reformulate the problem (P0) as

$$\begin{aligned} (\mathbf{P1.A}) \quad & \max_{\mathbf{V}} \quad \text{tr} \left(\mathbf{A}(\phi) \mathbf{P} \mathbf{P}^H \mathbf{A}^H(\phi) \hat{\chi}^{-1} \right) \\ \text{subject to } & (C.1) - (C.4) \text{ of } (\mathbf{P0}), \end{aligned} \quad (23)$$

Lemma 2: A lower bound for $\text{tr} \left(\mathbf{A}(\phi) \mathbf{P} \mathbf{P}^H \mathbf{A}^H(\phi) \hat{\chi}^{-1} \right)$ can be given as

$$\text{tr} \left(\mathbf{A}(\phi) \mathbf{P} \mathbf{P}^H \mathbf{A}^H(\phi) \hat{\chi}^{-1} \right) \geq \frac{\varphi R^2}{I^{RAD} + M \sigma_R^2}, \quad (24)$$

where $\varphi = \text{tr}(\mathbf{P} \mathbf{P}^H)$ and I^{RAD} is the total interference power from the CS to the MIMO RS, given as

$$\begin{aligned} I^{RAD} = & \sum_{k=1}^K \text{tr} \left\{ \mathbf{G}_{RU_k} \left(\mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H + \psi \text{diag} \left(\mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H \right) \right) (\mathbf{G}_{RU_k})^H \right\} \\ & + \sum_{j=1}^J \text{tr} \left\{ \mathbf{G}_{RB} \left(\mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H + \psi \text{diag} \left(\mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H \right) \right) (\mathbf{G}_{RB})^H \right\}, \end{aligned} \quad (25)$$

Proof: The proof is given in Appendix B. ■

From Lemma 2, the problem (P1.A) can be equivalently transformed into a interference minimisation problem as

$$\begin{aligned} (\mathbf{P1.B}) \quad & \min_{\mathbf{V}} \quad I^{RAD} \\ \text{subject to } & (C.1) - (C.4) \text{ of } (\mathbf{P0}). \end{aligned} \quad (26)$$

Now, for analytical simplicity, we write the QoS constraints (C.1) and (C.2) in terms of mean squared errors (MSE). To calculate the MSE, we first apply linear receive filters $\mathbf{U}_k^{UL} \in \mathbb{C}^{N_0 \times d_k^{UL}}$ and $\mathbf{U}_j^{DL} \in \mathbb{C}^{N_j \times d_j^{DL}}$ to \mathbf{y}_0 and \mathbf{y}_j^{DL} to obtain the source symbols. Accordingly, the extracted k -th UL user's symbols at the BS and j -th DL user's symbol from the BS can be given by

$$\begin{aligned} \hat{\mathbf{s}}_k^{UL} = & (\mathbf{U}_k^{UL})^H \left(\sum_{k=1}^K \mathbf{H}_k^{UL} (\mathbf{V}_k^{UL} \mathbf{s}_k^{UL} + \mathbf{c}_k^{UL}) + \mathbf{H}_0 \left(\sum_{j=1}^J \mathbf{V}_j^{DL} \mathbf{s}_j^{DL} + \mathbf{c}_0 \right) + \mathbf{e}_0 + \mathbf{W}_{BR}^{DL} \mathbf{s}_R + \mathbf{n}_0 \right), \end{aligned} \quad (27)$$

$$\begin{aligned} \hat{\mathbf{s}}_j^{DL} = & (\mathbf{U}_j^{DL})^H \left(\mathbf{H}_j^{DL} \left(\sum_{j=1}^J \mathbf{V}_j^{DL} \mathbf{s}_j^{DL} + \mathbf{c}_0 \right) + \sum_{k=1}^K \mathbf{H}_{jk}^{DU} (\mathbf{V}_k^{UL} \mathbf{s}_k^{UL} + \mathbf{c}_k^{UL}) + \mathbf{e}_j^{DL} + \mathbf{W}_j^{DL} \mathbf{s}_R + \mathbf{n}_j^{DL} \right). \end{aligned} \quad (28)$$

Hence, the MSE for k -th UL and j -th DL users can be respectively given as

$$\mathbf{E}_k^{UL}(\{\mathbf{U}\}, \{\mathbf{V}\}) = \left((\mathbf{U}_k^{UL})^H \mathbf{H}_k^{UL} \mathbf{V}_k^{UL} - \mathbf{I}_{d_k^{UL}} \right) \quad (29)$$

$$\times \left((\mathbf{U}_k^{UL})^H \mathbf{H}_k^{UL} \mathbf{V}_k^{UL} - \mathbf{I}_{d_k^{UL}} \right)^H + (\mathbf{U}_k^{UL})^H \mathbf{\Sigma}_k^{UL} \mathbf{U}_k^{UL},$$

$$\mathbf{E}_j^{DL}(\{\mathbf{U}\}, \{\mathbf{V}\}) = \left((\mathbf{U}_j^{DL})^H \mathbf{H}_j^{DL} \mathbf{V}_j^{DL} - \mathbf{I}_{d_j^{DL}} \right) \quad (30)$$

$$\times \left((\mathbf{U}_j^{DL})^H \mathbf{H}_j^{DL} \mathbf{V}_j^{DL} - \mathbf{I}_{d_j^{DL}} \right)^H + (\mathbf{U}_j^{DL})^H \mathbf{\Sigma}_j^{DL} \mathbf{U}_j^{DL},$$

where $\mathbf{U} = \{\mathbf{U}_k^{UL}, \mathbf{U}_j^{DL}\}$ denotes the set of all receive beamforming matrices. Note that, for fixed transmit beamforming

⁵Note that approximation of $\mathbf{\Sigma}_k^{UL}$ and $\mathbf{\Sigma}_j^{DL}$ is a practical assumption [24]. The values of ψ and v are much lower than 1. However, their values might not be negligible under a strong SI channel [24].

$$\begin{aligned}\Sigma_k^{UL} &\approx \sum_{j \neq k}^K \mathbf{H}_j^{UL} \mathbf{V}_j^{UL} (\mathbf{V}_j^{UL})^H (\mathbf{H}_j^{UL})^H + \psi \sum_{j=1}^K \mathbf{H}_j^{UL} \text{diag} \left(\mathbf{V}_j^{UL} (\mathbf{V}_j^{UL})^H \right) (\mathbf{H}_j^{UL})^H \\ &\quad + \sum_{j=1}^J \mathbf{H}_0 \left(\mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H + \psi \text{diag} \left(\mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H \right) \right) \mathbf{H}_0^H + P_R \left(\mathbf{W}_{BR}^{DL} (\mathbf{W}_{BR}^{DL})^H \right) + \sigma_0^2 \mathbf{I}_{N_0} \\ &\quad + v \sum_{j=1}^J \text{diag} \left(\mathbf{H}_0 \mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H \mathbf{H}_0^H \right) + v \sum_{j=1}^K \text{diag} \left(\mathbf{H}_j^{UL} \mathbf{V}_j^{UL} (\mathbf{V}_j^{UL})^H (\mathbf{H}_j^{UL})^H \right),\end{aligned}\quad (18)$$

$$\begin{aligned}\Sigma_j^{DL} &\approx \sum_{i \neq j}^J \mathbf{H}_i^{DL} \mathbf{V}_i^{DL} (\mathbf{V}_i^{DL})^H (\mathbf{H}_i^{DL})^H + \psi \sum_{i=1}^J \mathbf{H}_i^{DL} \text{diag} \left(\mathbf{V}_i^{DL} (\mathbf{V}_i^{DL})^H \right) (\mathbf{H}_i^{DL})^H \\ &\quad + \sum_{k=1}^K \mathbf{H}_{jk}^{DU} \left(\mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H + \psi \text{diag} \left(\mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H \right) \right) (\mathbf{H}_{jk}^{DU})^H + P_R \left(\mathbf{W}_j^{DL} (\mathbf{W}_j^{DL})^H \right) + \sigma_j^2 \mathbf{I}_{N_j} \\ &\quad + v \sum_{k=1}^K \text{diag} \left(\mathbf{H}_{jk}^{DU} \mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H (\mathbf{H}_{jk}^{DU})^H \right) + v \sum_{i=1}^J \text{diag} \left(\mathbf{H}_j^{DL} \mathbf{V}_i^{DL} (\mathbf{V}_i^{DL})^H (\mathbf{H}_j^{DL})^H \right).\end{aligned}\quad (19)$$

TABLE I: Simplification of Notations

\mathbf{H}_{ij}	$\mathbf{H}_j^{UL}, i \in \mathcal{S}^{UL}, j \in \mathcal{S}^{UL}; \mathbf{H}_0, i \in \mathcal{S}^{UL}, j \in \mathcal{S}^{DL}; \mathbf{H}_{ij}^{DU}, i \in \mathcal{S}^{DL}, j \in \mathcal{S}^{UL}; \mathbf{H}_i^{DL}, i \in \mathcal{S}^{DL}, j \in \mathcal{S}^{DL}$
\mathbf{G}_{R_j}	$\mathbf{G}_{RU_j}, j \in \mathcal{S}^{UL}; \mathbf{G}_{RB}, j \in \mathcal{S}^{DL}$
\mathbf{W}_i	$\mathbf{W}_i^{DL}, i \in \mathcal{S}^{DL}; \mathbf{W}_{BR}^{DL}, i \in \mathcal{S}_{BR}^{DL}$
\mathbf{n}_i	$\mathbf{n}_0, i \in \mathcal{S}^{UL}; \mathbf{n}_i^{DL}, i \in \mathcal{S}^{DL}$
$\tilde{N}_i(\tilde{M}_i)$	$N_0(M_i), i \in \mathcal{S}^{UL}; N_i(M_0), i \in \mathcal{S}^{DL}$

matrices, the optimal receive beamforming matrix at the BS for the k -th UL user and at the j -th DL user are MMSE receivers, which can be expressed as (31) and (32) shown on the top of the next page. Accordingly, substituting (31) and (32) in (29) and (30) and using the property $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}$ [30], the MSE matrices for the k -th UL and j -th DL users can be written as

$$\begin{aligned}\mathbf{E}_{k,\text{MMSE}}^{UL}(\{\mathbf{V}\}) &\quad (33) \\ &= \left(\mathbf{I}_k^{UL} + (\mathbf{V}_k^{UL})^H (\mathbf{H}_k^{UL})^H (\Sigma_k^{UL})^{-1} \mathbf{H}_k^{UL} \mathbf{V}_k^{UL} \right)^{-1},\end{aligned}$$

$$\begin{aligned}\mathbf{E}_{j,\text{MMSE}}^{DL}(\{\mathbf{V}\}) &\quad (34) \\ &= \left(\mathbf{I}_j^{DL} + (\mathbf{V}_j^{DL})^H (\mathbf{H}_j^{DL})^H (\Sigma_j^{DL})^{-1} \mathbf{H}_j^{DL} \mathbf{V}_j^{DL} \right)^{-1}.\end{aligned}$$

From (33)–(34) and (20)–(21), we obtain the relation between rate and MSE as

$$R_k^{UL} = \log_2 \left| (\mathbf{E}_{\text{MMSE},k}^{UL}(\{\mathbf{V}\}))^{-1} \right|; \quad (35)$$

$$R_j^{DL} = \log_2 \left| (\mathbf{E}_{\text{MMSE},j}^{DL}(\{\mathbf{V}\}))^{-1} \right|. \quad (36)$$

Before proceeding to the next section, we simplify the notations by combining UL and DL channels similar to [18]. Let the symbols \mathcal{S}^{UL} and \mathcal{S}^{DL} denote the set of K UL and J DL channels, respectively, and the channels from RS to BS and the set of J DL channels from RS to DL users are denoted by \mathcal{S}_{BU}^{DL} and \mathcal{S}_{BR}^{DL} , respectively. Now, denoting \mathbf{V}_i^X , \mathbf{U}_i^X , d_i^X and Σ_i^X , $X \in \{UL, DL\}$ as \mathbf{V}_i , \mathbf{U}_i , d_i and Σ_i , respectively, and expressing \mathbf{H}_{ij} , \mathbf{G}_{R_j} , \mathbf{n}_i , and receive (transmit) antenna numbers $\tilde{N}_i(\tilde{M}_i)$ as shown in Table I, the MSE of the i -th link, $i \in \mathcal{S} \triangleq \mathcal{S}^{UL} \cup \mathcal{S}^{DL}$ and the interference power at the MIMO RS, I^{RAD} can be rewritten, respectively as

$$\mathbf{E}_i = (\mathbf{U}_i^H \mathbf{H}_{ii} \mathbf{V}_i - \mathbf{I}_{d_i}) (\mathbf{U}_i^H \mathbf{H}_{ii} \mathbf{V}_i - \mathbf{I}_{d_i})^H + \mathbf{U}_i^H \Sigma_i \mathbf{U}_i, \quad (37)$$

where

$$\begin{aligned}\Sigma_i &= \sum_{j \in \mathcal{S}, j \neq i} \mathbf{H}_{ij} \mathbf{V}_j \mathbf{V}_j^H \mathbf{H}_{ij}^H + \psi \sum_{j \in \mathcal{S}} \mathbf{H}_{ij} \text{diag} \left(\mathbf{V}_j \mathbf{V}_j^H \right) \mathbf{H}_{ij}^H \\ &\quad + v \sum_{j \in \mathcal{S}} \text{diag} \left(\mathbf{H}_{ij} \mathbf{V}_j \mathbf{V}_j^H \mathbf{H}_{ij}^H \right) + P_R \left(\mathbf{W}_i (\mathbf{W}_i)^H \right) \\ &\quad + \sigma_i^2 \mathbf{I}_{\tilde{N}_i},\end{aligned}\quad (38)$$

$$I^{RAD} = \sum_{j \in \mathcal{S}} \text{tr} \left\{ \mathbf{G}_{R_j} (\mathbf{V}_j \mathbf{V}_j^H + \psi \text{diag} \left(\mathbf{V}_j \mathbf{V}_j^H \right)) \mathbf{G}_{R_j}^H \right\}. \quad (39)$$

Now, using (35), (36), the above simplified notations, and epigraph method, [31] the optimisation problem **(P1.B)** can be equivalently reformulated as

$$\textbf{(P1.C)} \quad \min_{\mathbf{V}, \Gamma} \quad \Gamma \quad (40)$$

$$\text{subject to} \quad (C.1) \quad I^{RAD} \leq \Gamma$$

$$(C.2) \quad \log_2 \left| (\mathbf{E}_{i,\text{MMSE}}(\{\mathbf{V}_i\}))^{-1} \right| \geq R_{i,\min}^{UL}, \quad i \in \mathcal{S}^{UL},$$

$$(C.3) \quad \log_2 \left| (\mathbf{E}_{i,\text{MMSE}}(\{\mathbf{V}_i\}))^{-1} \right| \geq R_{i,\min}^{DL}, \quad i \in \mathcal{S}^{DL},$$

$$(C.4) \quad \text{tr} \left\{ \mathbf{V}_i \mathbf{V}_i^H \right\} \leq P_i, \quad i \in \mathcal{S}^{UL},$$

$$(C.5) \quad \sum_{i \in \mathcal{S}^{DL}} \text{tr} \left\{ \mathbf{V}_i \mathbf{V}_i^H \right\} \leq P_0,$$

V. ROBUST BEAMFORMER DESIGN AT FD MIMO CS

In order to design a more practical and robust system, in this section, we assume that perfect CSI knowledge of all the concerned channels is unavailable at the CS⁶. By considering the worst-case (norm-bounded error) model [32], the channel uncertainties can be constructed as

$$\mathbf{H}_{ij} \in \mathcal{H}_{ij} = \left\{ \tilde{\mathbf{H}}_{ij} + \Delta_{ij} : \|\Delta_{ij}\|_F \leq \delta_{ij} \right\}, \quad (41)$$

$$\mathbf{G}_{R_j} \in \mathcal{G}_j = \left\{ \tilde{\mathbf{G}}_{R_j} + \Lambda : \|\Lambda\|_F \leq \theta, j \in \mathcal{S} \right\}, \quad (42)$$

where the nominal values of the CSI are denoted by $\tilde{\mathbf{H}}_{ij}$ and $\tilde{\mathbf{G}}_j$, while Δ_{ij} and Λ_j represent the channel error matrix and δ_{ij} and θ_j express the uncertainty bounds.

Lemma 3: Assume that $\mathbf{F} \in \mathbb{C}^{d \times d}$ is a positive semi-definite matrix and $|\mathbf{F}| \leq 1$. Then the maximisation of the function

⁶Note that this assumption doesn't obtrude with the assumption that RS has full CSI knowledge of the CS's channels and adheres to Section II.3.

$$\mathbf{U}_k^{UL*} = \arg \min_{\mathbf{U}_k^{UL}} \text{tr}(\mathbf{E}_k^{UL}) = (\mathbf{V}_k^{UL})^H (\mathbf{H}_k^{UL})^H \left(\mathbf{H}_k^{UL} \mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H (\mathbf{H}_k^{UL})^H + \Sigma_k^{UL} \right)^{-1}; \quad (31)$$

$$\mathbf{U}_j^{DL*} = \arg \min_{\mathbf{U}_j^{DL}} \text{tr}(\mathbf{E}_j^{DL}) = (\mathbf{V}_j^{DL})^H (\mathbf{H}_j^{DL})^H \left(\mathbf{H}_j^{DL} \mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H (\mathbf{H}_j^{DL})^H + \Sigma_j^{DL} \right)^{-1}. \quad (32)$$

$f(\mathbf{Q}) = -\text{tr}(\mathbf{Q}\mathbf{F}) + \log|\mathbf{Q}| + d$ is equivalent to $\log|\mathbf{F}^{-1}|$, i.e.,

$$\max_{\mathbf{Q} \in \mathbb{C}^{d \times d}, \mathbf{Q} \succ 0} f(\mathbf{Q}) = \log|\mathbf{F}^{-1}|, \quad (43)$$

where the optimum value $\mathbf{Q}^{\text{opt}} = \mathbf{F}^{-1}$.

Now, applying Lemma 3 and decomposing $\mathbf{Q} = \mathbf{B}_i \mathbf{B}_i^H$, the problem (P1.C) under channel uncertainties can be expressed as

$$(\mathbf{P2}) \quad \min_{\mathbf{V}, \Gamma} \quad \Gamma \quad (44)$$

subject to (C.1) $I^{RAD} \leq \Gamma$, $\forall \mathbf{G}_{R_j} \in \mathcal{G}_j$,

$$(C.2) \quad \min_{\Delta} \max_{\{\mathbf{U}_i, \mathbf{B}_i\}} -\text{tr}(\mathbf{B}_i^H \mathbf{E}_i \mathbf{B}_i) + 2 \log|\mathbf{B}_i| + d_i \\ \geq \log(2) R_{i,min}^{UL}, \quad i \in \mathcal{S}^{UL},$$

$$(C.3) \quad \min_{\Delta} \max_{\{\mathbf{U}_i, \mathbf{B}_i\}} -\text{tr}(\mathbf{B}_i^H \mathbf{E}_i \mathbf{B}_i) + 2 \log|\mathbf{B}_i| + d_i \\ \geq \log(2) R_{i,min}^{DL}, \quad i \in \mathcal{S}^{DL},$$

$$(C.4) \quad \text{tr}\{\mathbf{V}_i \mathbf{V}_i^H\} \leq P_i, \quad i \in \mathcal{S}^{UL},$$

$$(C.5) \quad \sum_{i \in \mathcal{S}^{DL}} \text{tr}\{\mathbf{V}_i \mathbf{V}_i^H\} \leq P_0,$$

where $\Delta = \{\Delta_{ij} : \forall (i, j)\}$ and $\mathbf{B}_i \in \mathbb{C}^{d_i \times d_i}$, $i \in \mathcal{S}$, is a weight matrix. Due to the constraint (C.1) in (44) and the inner maximization in (C.2) and (C.3), the problem (P2) is intractable. To make the problem tractable, we consider the following max-min inequality [33]

$$\min_{\Delta, \Lambda} \max_{\{\mathbf{U}_i, \mathbf{Q}_i\}} (\cdot) \geq \max_{\{\mathbf{U}_i, \mathbf{Q}_i\}} \min_{\Delta, \Lambda} (\cdot) \geq \log(2) R_{i,min}, i \in \mathcal{S}. \quad (45)$$

Now, we convert $\text{tr}(\mathbf{B}_i^H \mathbf{E}_i \mathbf{B}_i)$ and I^{RAD} into vector forms, given in the following lemma.

Lemma 4: $\text{tr}(\mathbf{B}_i^H \mathbf{E}_i \mathbf{B}_i)$ and I^{RAD} can be written in the form of vectors as $\text{tr}(\mathbf{B}_i^H \mathbf{E}_i \mathbf{B}_i) = \|\mathbf{z}_{ij}\|_2^2$ and $I^{RAD} = \|\boldsymbol{\iota}\|_2^2$, where \mathbf{z}_{ij} and $\boldsymbol{\iota}$ are defined as⁷

$$\mathbf{z}_{ij} = \begin{bmatrix} \text{vec}(\mathbf{B}_i^H (\mathbf{U}_i^H \mathbf{H}_{ii} \mathbf{V}_i - \mathbf{I}_{d_i})) \\ [(\mathbf{V}_j^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H)) \text{vec}(\mathbf{H}_{ij})]_{j \in \mathcal{S}, j \neq i} \\ \left[\left[\sqrt{\psi} ((\Xi_\ell \mathbf{V}_j)^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H)) \text{vec}(\mathbf{H}_{ij}) \right]_{\ell \in \mathcal{D}_j^{(T)}} \right]_{j \in \mathcal{S}} \\ \left[\left[\sqrt{v} (\mathbf{V}_j^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H \Xi_\ell)) \text{vec}(\mathbf{H}_{ij}) \right]_{\ell \in \mathcal{D}_i^{(R)}} \right]_{j \in \mathcal{S}} \\ P_R (\mathbf{I}_R^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H)) \text{vec}(\mathbf{W}_i) \\ \sigma_i \text{vec}(\mathbf{B}_i^H \mathbf{U}_i^H) \end{bmatrix},$$

$$\boldsymbol{\iota} = \begin{bmatrix} [(\mathbf{V}_j^T \otimes \mathbf{I}_T) \text{vec}(\mathbf{G}_j)]_{j \in \mathcal{S}} \\ \sqrt{\psi} \left[\left[((\Xi_\ell \mathbf{V}_j)^T \otimes \mathbf{I}_T) \text{vec}(\mathbf{G}_j) \right]_{\ell \in \mathcal{D}_j^{(T)}} \right]_{j \in \mathcal{S}} \end{bmatrix}. \quad (46)$$

Here, Ξ_ℓ represents a square matrix with zero as elements, except for the ℓ -th diagonal element, which is equal to 1, and $\mathcal{D}_j^{(R)}$ and $\mathcal{D}_j^{(T)}$ denote the set $\{1 \cdots \tilde{N}_j\}$ and $\{1 \cdots \tilde{M}_j\}$, respectively.

Proof: The proof is provided in Appendix C. ■

From Lemma 4, the constraint (C.2) of (P2) can be rewritten as

$$\min_{\Delta} \max_{\{\mathbf{U}_i, \mathbf{B}_i\}} -\text{tr}(\mathbf{B}_i^H \mathbf{E}_i \mathbf{B}_i) + 2 \log|\mathbf{B}_i| + d_i \\ \geq \log(2) R_{i,min}^{UL}, \quad i \in \mathcal{S}^{UL}, \quad (47)$$

$$\Rightarrow \begin{cases} (C.2a) & d_i + 2 \log|\mathbf{B}_i| - \sum_{j \in \mathcal{S}} \lambda_{ij} \\ & \geq \log(2) R_{i,min}^{UL}, \quad \forall i \in \mathcal{S}^{UL} \\ (C.2b) & \max_{\Delta} \|\hat{\mathbf{z}}_{ij} + \hat{\mathbf{Z}}_{ij} \text{vec}(\Delta_{ij})\|_2^2 \leq \lambda_{ij}, \\ & \|\Delta_{ij}\|_F \leq \delta_{ij}, \quad \forall i, j \in \mathcal{S}^{UL}, \end{cases} \quad (48)$$

where $\hat{\mathbf{z}}_{ij}$ and $\hat{\mathbf{Z}}_{ij}$ are defined as

$$\hat{\mathbf{z}}_{ij} = \begin{bmatrix} \text{vec}(\mathbf{B}_i^H (\mathbf{U}_i^H \tilde{\mathbf{H}}_{ii} \mathbf{V}_i - \mathbf{I}_{d_i})) \\ [(\mathbf{V}_j^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H)) \text{vec}(\tilde{\mathbf{H}}_{ij})]_{j \in \mathcal{S}^{UL}, j \neq i} \\ \left[\left[\sqrt{\psi} ((\Xi_\ell \mathbf{V}_j)^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H)) \text{vec}(\tilde{\mathbf{H}}_{ij}) \right]_{\ell \in \mathcal{D}_j^{(T)}} \right]_{j \in \mathcal{S}} \\ \left[\left[\sqrt{v} (\mathbf{V}_j^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H \Xi_\ell)) \text{vec}(\tilde{\mathbf{H}}_{ij}) \right]_{\ell \in \mathcal{D}_i^{(R)}} \right]_{j \in \mathcal{S}} \\ P_R (\mathbf{I}_R^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H)) \text{vec}(\mathbf{W}_i) \\ \sigma_i \text{vec}(\mathbf{B}_i^H \mathbf{U}_i^H) \end{bmatrix} \quad (49)$$

$$\hat{\mathbf{Z}}_i = \begin{bmatrix} (\mathbf{V}_i^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H)) \\ [(\mathbf{V}_j^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H))]_{j \in \mathcal{S}^{UL}, j \neq i} \\ \left[\left[\sqrt{\psi} ((\Xi_\ell \mathbf{V}_j)^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H)) \right]_{\ell \in \mathcal{D}_j^{(T)}} \right]_{j \in \mathcal{S}} \\ \left[\left[\sqrt{v} (\mathbf{V}_j^T \otimes (\mathbf{B}_i^H \mathbf{U}_i^H \Xi_\ell)) \right]_{\ell \in \mathcal{D}_i^{(R)}} \right]_{j \in \mathcal{S}} \\ \mathbf{0}_{d_i \tilde{N}_i \times \tilde{N}_i \tilde{M}} \\ \mathbf{0}_{d_i \tilde{N}_i \times \tilde{N}_i \tilde{M}} \end{bmatrix}. \quad (50)$$

Now, using Lemma 5 from Appendix D, we relax the semi-infiniteness of the constraint (C.2b) in (56) and then apply Lemma 6 from Appendix D, to convert the norm constraint of (C.2b) into a LMI form as

$$(C.2b) \quad \begin{bmatrix} \lambda_{ij} & \hat{\mathbf{z}}_{ij}^H \\ \hat{\mathbf{z}}_{ij} & \mathbf{I}_{d_i d_j} \end{bmatrix} \\ + \begin{bmatrix} 0 & \text{vec}(\Delta_{ij})^H \hat{\mathbf{Z}}_{ij}^H \\ \hat{\mathbf{Z}}_{ij} \text{vec}(\Delta_{ij}) & \mathbf{0}_{d_i d_j \times d_i d_j} \end{bmatrix} \succeq 0. \quad (51)$$

⁷For sake of simplicity, we consider $\tilde{M} = M_0 = M_i$, $i \in \mathcal{S}^{UL}$.

By choosing

$$\mathbf{A} = \begin{bmatrix} \lambda_{ij} & \hat{\mathbf{z}}_{ij}^H \\ \hat{\mathbf{z}}_{ij} & \mathbf{I}_{d_i d_j} \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} \mathbf{0}_{\tilde{N}_i \tilde{M} \times 1} & \hat{\mathbf{z}}_{ij}^H \\ \mathbf{0}_{\tilde{N}_i \tilde{M} \times 1} & \mathbf{I}_{d_i d_j} \end{bmatrix},$$

$$\mathbf{X} = \text{vec}(\Delta_{ij}), \quad \mathbf{Q} = [-1, \mathbf{0}_{1 \times d_{ij}}], \quad (52)$$

and applying Lemma 5, the constraint (C.2b) of (56) is equivalently rewritten as

$$\begin{bmatrix} \lambda_{ij} - \epsilon_{ij} & \hat{\mathbf{z}}_{ij}^H & \mathbf{0}_{1 \times \tilde{N}_i \tilde{M}} \\ \hat{\mathbf{z}}_{ij} & \mathbf{I}_{d_i d_j} & -\delta_{ij} \hat{\mathbf{z}}_{ij}^H \\ \mathbf{0}_{\tilde{N}_i \tilde{M} \times 1} & -\delta_{ij} \hat{\mathbf{z}}_{ij}^H & \epsilon_{ij} \mathbf{I}_{\tilde{N}_i \tilde{M}} \end{bmatrix} \succeq 0, \quad \forall i, j \in \mathcal{S}^{UL}, \quad (53)$$

$$\epsilon_{ij} \geq 0, \quad \forall i, j \in \mathcal{S}^{UL}, \quad (54)$$

where $\epsilon = \{\epsilon_{ij} : \forall(i, j)\}$. Similar to the transformation of the constraint (C.2) of (P2), the constraint (C.3) of (P2) can be equivalently rewritten as

$$\min_{\Delta} \max_{\{\mathbf{U}_i, \mathbf{B}_i\}} -\text{tr}(\mathbf{B}_i^H \mathbf{E}_i \mathbf{B}_i) + 2 \log |\mathbf{B}_i| + d_i$$

$$\geq \log(2) R_{i, \min}^{DL}, \quad i \in \mathcal{S}^{DL}, \quad (55)$$

$$\Rightarrow \begin{cases} (C.3a) & d_i + 2 \log |\mathbf{B}_i| - \sum_{j \in \mathcal{S}} \lambda_{ij} \\ & \geq \log(2) R_{i, \min}^{DL}, \quad \forall i \in \mathcal{S}^{DL} \\ (C.3b) & \max_{\Delta} \|\hat{\mathbf{z}}_{ij} + \hat{\mathbf{z}}_{ij} \text{vec}(\Delta_{ij})\|_2^2 \leq \lambda_{ij}, \\ & \|\Delta_{ij}\|_F \leq \delta_{ij}, \quad \forall i, j \in \mathcal{S}^{DL}, \end{cases} \quad (56)$$

where (C.3b) is expressed as

$$\begin{bmatrix} \lambda_{ij} - \epsilon_{ij} & \hat{\mathbf{z}}_{ij}^H & \mathbf{0}_{1 \times \tilde{N}_i \tilde{M}} \\ \hat{\mathbf{z}}_{ij} & \mathbf{I}_{d_i d_j} & -\delta_{ij} \hat{\mathbf{z}}_{ij}^H \\ \mathbf{0}_{\tilde{N}_i \tilde{M} \times 1} & -\delta_{ij} \hat{\mathbf{z}}_{ij}^H & \epsilon_{ij} \mathbf{I}_{\tilde{N}_i \tilde{M}} \end{bmatrix} \succeq 0, \quad \forall i \in \mathcal{S}^{DL}, j \in \mathcal{S}, \quad (57)$$

$$\epsilon_{ij} \geq 0, \quad \forall i \in \mathcal{S}^{DL}, j \in \mathcal{S}. \quad (58)$$

In similar way, we also transform the constraint (C.1) of (P2) as

$$\begin{bmatrix} \Gamma & \tilde{\mathbf{t}}^H \\ \tilde{\mathbf{t}} & \mathbf{I}_B \end{bmatrix} + \begin{bmatrix} 0 & \boldsymbol{\iota}_\Lambda^H \\ \boldsymbol{\iota}_\Lambda & \mathbf{0}_{B \times B} \end{bmatrix} \succeq 0, \quad (59)$$

where $B = R \sum_{j \in \mathcal{S}} (d_j + \tilde{M}_j)$,

$$\tilde{\mathbf{t}} = \begin{bmatrix} [(\mathbf{V}_j^T \otimes \mathbf{I}_T) \text{vec}(\tilde{\mathbf{G}}_{lj})]_{j \in \mathcal{S}} \\ \sqrt{\psi} \left[[((\Xi_\ell \mathbf{V}_j)^T \otimes \mathbf{I}_T) \text{vec}(\tilde{\mathbf{G}}_{lj})]_{\ell \in \mathcal{D}_j^{(T)}} \right]_{j \in \mathcal{S}} \end{bmatrix},$$

$$\boldsymbol{\iota}_\Lambda = \underbrace{\begin{bmatrix} [(\mathbf{V}_j^T \otimes \mathbf{I}_R)]_{j \in \mathcal{S}} \\ \sqrt{\psi} \left[[((\Xi_\ell \mathbf{V}_j)^T \otimes \mathbf{I}_R)]_{\ell \in \mathcal{D}_j^{(T)}} \right]_{j \in \mathcal{S}} \end{bmatrix}}_{\mathbf{E}_\Lambda} \text{vec}(\Lambda). \quad (60)$$

Finally, the constraint (C.1) of (P2) is equivalently rewritten as

$$\begin{bmatrix} \Gamma - \eta & \tilde{\mathbf{t}}^H & \mathbf{0}_{1 \times R\tilde{M}} \\ \tilde{\mathbf{t}} & \mathbf{I}_B & -\theta \mathbf{E}_\Lambda \\ \mathbf{0}_{R\tilde{M} \times 1} & -\theta \mathbf{E}_\Lambda^H & \eta \mathbf{I}_{R\tilde{M}} \end{bmatrix} \succeq 0, \quad (61)$$

$$\eta \geq 0. \quad (62)$$

Using the relaxed LMIs in (53), (57) and (61), the problem (P2) can be written as a SDP problem, given as

$$(\mathbf{P3}) \quad \min_{\mathbf{V}, \Gamma, \mathbf{U}, \mathbf{B}, \boldsymbol{\lambda}, \epsilon \geq 0, \eta \geq 0} \Gamma \quad (63)$$

$$\text{subject to (C.1)} \quad \begin{bmatrix} \Gamma - \eta & \tilde{\mathbf{t}}^H & \mathbf{0}_{1 \times R\tilde{M}} \\ \tilde{\mathbf{t}} & \mathbf{I}_B & -\theta \mathbf{E}_\Lambda \\ \mathbf{0}_{R\tilde{M} \times 1} & -\theta \mathbf{E}_\Lambda^H & \eta \mathbf{I}_{R\tilde{M}} \end{bmatrix} \succeq 0,$$

$$(C.2a) \quad d_i + 2 \log |\mathbf{B}_i| - \sum_{j \in \mathcal{S}} \lambda_{ij} \geq \log(2) R_{i, \min}^{UL}, \quad \forall i \in \mathcal{S}^{UL},$$

$$(C.2b) \quad \begin{bmatrix} \lambda_{ij} - \epsilon_{ij} & \hat{\mathbf{z}}_{ij}^H & \mathbf{0}_{1 \times \tilde{N}_i \tilde{M}} \\ \hat{\mathbf{z}}_{ij} & \mathbf{I}_{d_i d_j} & -\delta_{ij} \hat{\mathbf{z}}_{ij}^H \\ \mathbf{0}_{\tilde{N}_i \tilde{M} \times 1} & -\delta_{ij} \hat{\mathbf{z}}_{ij}^H & \epsilon_{ij} \mathbf{I}_{\tilde{N}_i \tilde{M}} \end{bmatrix} \succeq 0,$$

$$\forall i \in \mathcal{S}^{UL}, j \in \mathcal{S},$$

$$(C.3a) \quad d_i + 2 \log |\mathbf{B}_i| - \sum_{j \in \mathcal{S}} \lambda_{ij} \geq \log(2) R_{i, \min}^{DL}, \quad \forall i \in \mathcal{S}^{DL},$$

$$(C.3b) \quad \begin{bmatrix} \lambda_{ij} - \epsilon_{ij} & \hat{\mathbf{z}}_{ij}^H & \mathbf{0}_{1 \times \tilde{N}_i \tilde{M}} \\ \hat{\mathbf{z}}_{ij} & \mathbf{I}_{d_i d_j} & -\delta_{ij} \hat{\mathbf{z}}_{ij}^H \\ \mathbf{0}_{\tilde{N}_i \tilde{M} \times 1} & -\delta_{ij} \hat{\mathbf{z}}_{ij}^H & \epsilon_{ij} \mathbf{I}_{\tilde{N}_i \tilde{M}} \end{bmatrix} \succeq 0,$$

$$\forall i \in \mathcal{S}^{DL}, j \in \mathcal{S},$$

$$(C.4) \quad \|\text{vec}(\mathbf{V}_i)\|_2^2 \leq P_i, \quad i \in \mathcal{S}^{UL},$$

$$(C.5) \quad \|\llbracket \text{vec}(\mathbf{V}_i) \rrbracket_{i \in \mathcal{S}^{DL}}\|_2^2 \leq P_0,$$

where $\mathbf{B} = \{\mathbf{B}_i, \forall i \in \mathcal{S}\}$ and $\boldsymbol{\lambda} = \{\lambda_{ij}, \forall i, j \in \mathcal{S}\}$.

Note that the optimisation problem (P3) is not jointly convex over the optimisation variables \mathbf{V} , \mathbf{U} and \mathbf{B} . However, it is separately convex over each of the variables. Therefore, we adopt an alternating algorithm to solve the problem. This alternating minimisation process is continued until a stationary point is obtained, or a pre-defined number of iterations is reached. In the following section, we provide details on the spectrum sharing algorithm, including the alternating optimisation of the above SDP problem.

VI. SPECTRUM SHARING ALGORITHM

In this section we summarise the roles played in spectrum sharing by the RS and CS in the form of algorithms. While Algorithm 1 presents the role played by the RS, Algorithm 2 illustrates the role of the CS in the proposed 2 tier spectrum sharing framework⁸. In particular, Algorithm 1 performs NSP towards the interference channels of the CS, thereby cancelling the interference from RS to CS. Alternatively, Algorithm 2 produces the optimal beamforming matrices at the transmitters and receivers of the CS, to minimise the interference from the CS to RS (equivalently maximises the PoD of RS), while maintaining a particular QoS for the cellular users. Note that Algorithm 2 is iterative in nature and solves a SDP problem in each iteration, which makes it computationally intensive. Below we provide some qualitative analysis on the complexity of Algorithm 2.

⁸Note that both algorithms are processed within the same coherence time interval.

Algorithm 1: Spectrum Sharing Phase at RS
I. Phase 1 [Initial Phase]:
A: Obtain CSI of $\{\mathbf{W}_i^{DL}, \forall i \in \mathcal{S}^{DL}, \mathbf{W}_{BR}^{DL}\} \subseteq \mathbf{W}$ through feedback from CS.
II. Phase 2 [Null-space Projection Phase]:
A: Perform SVD of \mathbf{W} , which is obtained from Phase 1.
B: Construct $\tilde{\mathbf{\Omega}}$ and $\hat{\mathbf{\Omega}}$ according to (13) and (14)
C: Design the projection matrix \mathbf{P} based on (15).
C: Output: Perform NSP by transmitting waveform $\mathbf{x}_R = \mathbf{P}\mathbf{s}_R$.

Algorithm 2: Spectrum Sharing Phase at CS
I. Phase 1 [Initial Phase]:
A: Obtain partial CSI of $\{\mathbf{G}_{RU_j}, \forall j \in \mathcal{S}^{UL}, \mathbf{G}_{RB}\}$.
B: Set minimum QoS requirements for UL and DL users: $R_{i,min}^{UL}$ and $R_{i,min}^{DL}$.
C: Initialize $\mathbf{V}^{[n]}$, $\mathbf{U}^{[n]}$ and $\mathbf{B}^{[n]}$.
D: Set iteration number $n = 0$, maximum iteration number $= n_{max}$.
II. Phase 2 [Beamforming Design Phase (Alternating Approach)]:
A: $n \leftarrow n + 1$. For fixed $\mathbf{U}_i^{[n-1]}$ and $\mathbf{B}_i^{[n-1]}$, update $\mathbf{V}^{[n]}$, $\forall i \in \mathcal{S}$ by solving problem (P3).
B: Update $\mathbf{B}_i^{[n]}$, $i \in \mathcal{S}$ by solving the problem (P3) for fixed $\mathbf{U}_i^{[n-1]}$ and $\mathbf{V}_i^{[n-1]}$.
C: Update $\mathbf{U}_i^{[n]}$, $i \in \mathcal{S}$ by solving the problem (P3) for fixed $\mathbf{V}_i^{[n-1]}$ and $\mathbf{B}_i^{[n-1]}$.
D: Repeat steps II.A – II.C until convergence or $n = n_{max}$.
E: Output: Optimal transceivers: $\{\mathbf{U}^*, \mathbf{V}^*\}$.

Computational complexity of Algorithm 2: The computational complexity mainly depends on the number of arithmetic operations required to process Phase 2 of Algorithm 2. In particular, a SDP problem is solved in Phase 2 in three steps, i.e., Step II.A to Step II.C. Hence, for comparison we first consider a standard real-valued SDP problem as

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{R}^n} \quad & \mathbf{c}^T \mathbf{x} \\ \text{subject to} \quad & \mathbf{A}_0 + \sum_{i=1}^n x_i \mathbf{A}_i \succeq \mathbf{0}, \text{ and } \|\mathbf{x}\|_2 \leq X, \end{aligned} \quad (65)$$

where \mathbf{A}_i is a symmetric block-diagonal matrix. If P is the diagonal block of matrix \mathbf{A}_i of size $a_l \times a_l$, $l = 1, \dots, P$, then the number of arithmetic operations required to solve (65) is upper-bounded by [34]

$$\mathcal{O}(1) \left(1 + \sum_{l=1}^P a_l \right)^{1/2} n \left(n^2 + n \sum_{l=1}^P a_l^2 + \sum_{l=1}^P a_l^3 \right). \quad (66)$$

Thus, using (66), we can compute the total number of arithmetic operations required to find the optimal \mathbf{V}_i , \mathbf{U}_i , and \mathbf{B}_i in Algorithm 2. For instance, in order to find \mathbf{V}_i , the number of diagonal blocks P is $|S^{UL}|(|S| + 1) + |S^{DL}||S| + 1$. The constraints (C.2b) and (C.3b) create the blocks of size $a_{ij} = 2(N_i \tilde{M} + d_i d_j + 1)$, $i \in \mathcal{S}^{DL}$, $j \in \mathcal{S}^{UL}$ and $a_{ij} = 2(\tilde{N}_i \tilde{M} + d_i d_j + 1)$, $i \in \mathcal{S}^{UL}$, $j \in \mathcal{S}^{DL}$, respectively. The size of blocks due to the constraint (C.1) is $a = 2(B + R\tilde{M} + 1)$, while the constraints for UL power in (C.4) and the BS power in (C.5) make the blocks of size $a_i = \tilde{M} d_i^{UL} + 1$, $i \in \mathcal{S}^{UL}$ and $a_i = \tilde{M} \sum_{i \in \mathcal{S}^{DL}} d_i^{DL} + 1$, respectively. Furthermore, the number required to compute the unknown variables is $n = \sum_{i \in \mathcal{S}} 2\tilde{M} d_i + 2|S| + 2$, where the term $\sum_{i \in \mathcal{S}} 2\tilde{M} d_i$ correlates with the real and image parts of \mathbf{V}_i and the remaining terms are due to the additional slack variables. Similarly, the number of arithmetic operations required for \mathbf{U}_i , $i \in \mathcal{S}$ and \mathbf{B}_i , $i \in \mathcal{S}$ can be calculated.

VII. NUMERICAL RESULTS

In this section, the performance of the proposed two-tier coexistence framework between a FD MU MIMO CS and a MIMO RS is analysed with the help of computer simulations⁹ under consideration of QoS of cellular users. The maximum number of iterations is set as 50 with a tolerance value of 10^{-4} . The initialisation points are selected using right singular matrices initialisation [35] and the results are averaged over 100 independent channel realisations.

1) *Simulation Setup:* To model the CS, we consider small cell deployments under the 3GPP LTE specifications. The motivation behind this are: 1) due to low transmit powers, short transmission distances and low mobility, small cells are considered suitable for implementation of FD technology [23], and 2) FCC has proposed the use of small cells in the 3.5 GHz band for spectrum sharing [2]. Accordingly, a single hexagonal cell of radius $r = 40$ m is considered, where the FD BS is located at the centre of the cell and a MIMO RS is located 400 m away from the circumference of the cell. The number of UL and DL users is set as $K = J = 2$ and each user, equipped with N antennas is randomly located in the cell. For simplicity, we consider $M_0 = N_0 = N = \tilde{N}$. Next, to model the path loss in the CS, we consider the close-in (CI) free space reference distance path loss model as given in [36]. The CI model is a generic model that describes the large-scale propagation path loss at all relevant frequencies (> 2 GHz). This model can be easily implemented in existing 3GPP models by replacing a floating constant with a frequency-dependent constant that represents free space path loss in the first meter of propagation and is given as $PL(f, d) = PL_F(f, d_0) + 10\alpha_c \log_{10}(d/d_0) + \mathcal{X}_\sigma$, $d > d_0$. Here, d_0 is a reference distance at which or closer to, the path loss inherits the characteristics of free-space path loss PL_F . Further, f is the carrier frequency, α_c is the path loss exponent, d is the distance between the transmitter and receiver and \mathcal{X}_σ is the shadow fading standard deviation. We consider $d_0 = 1$ m, $B = 100$ MHz, and carrier frequency $= 3.6$ GHz¹⁰.

The estimated channel gain between the BS and k th UL user can be described as $\tilde{\mathbf{H}}_k^{UL} = \sqrt{\varphi_k^{UL}} \hat{\mathbf{H}}_k^{UL}$, where $\hat{\mathbf{H}}_k^{UL}$ denotes small scale fading following a complex Gaussian distribution with zero mean and unit variance, and $\varphi_k^{UL} = 10^{(-A/10)}$, $A \in \{\text{LOS}, \text{NLOS}\}$ ¹¹ denotes the large scale fading consisting of path loss and shadowing. LOS and NLOS are computed based on a street canyon scenario [37]. The parameter α_c for LOS and NLOS are set as 2.0 and 3.1, respectively, while the value of shadow fading standard deviation σ for LOS and NLOS are 2.9 dB and 8.1 dB, respectively. Similarly, we define the channels between UL users and DL users, between BS and DL users, between BS and RS, and between UL

⁹For reference, the numerical results are obtained using MATLAB R2016b on a Linux server with Intel Xeon processor (16 cores, each clocked at 2 GHz) having 31.4 GiB of memory.

¹⁰The framework presented in this paper is not limited to any particular frequency band and can also be utilized in other spectrums proposed for sharing around the world, such as 2-4 GHz in the UK, 2.3-2.4 GHz in Europe, etc., albeit certain changes in frequency dependent path loss, line of sight propagation parameters, etc.

¹¹LOS=Line-of-sight and NLOS=Non-line-of-sight.

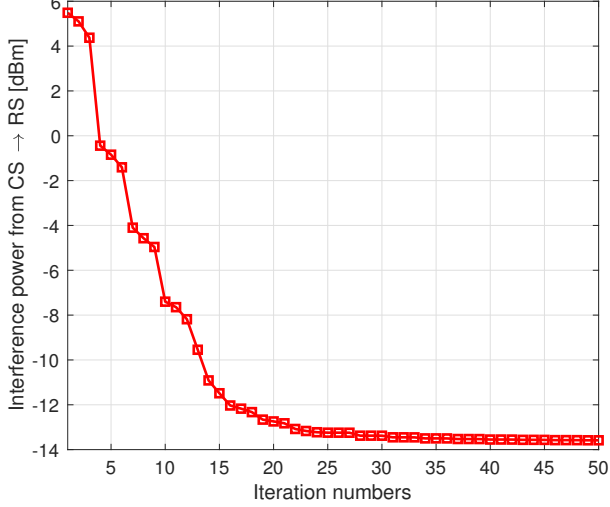


Fig. 2: Convergence of Algorithm 2.

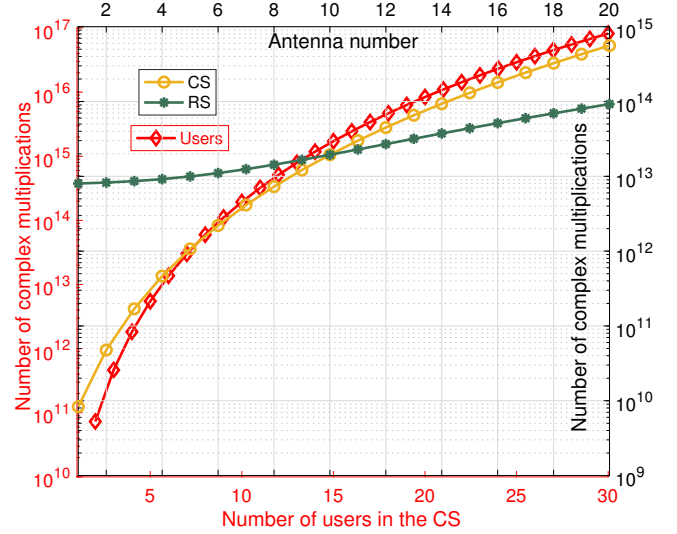


Fig. 3: Complexity of Algorithm 2.

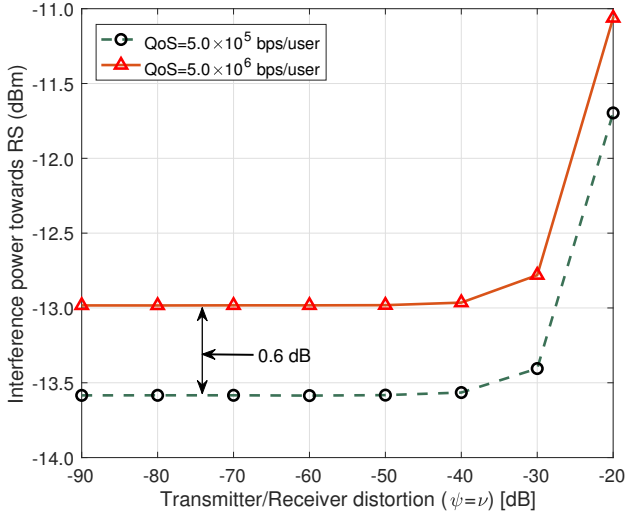


Fig. 4: Interference power towards RS vs. RSI at CS.

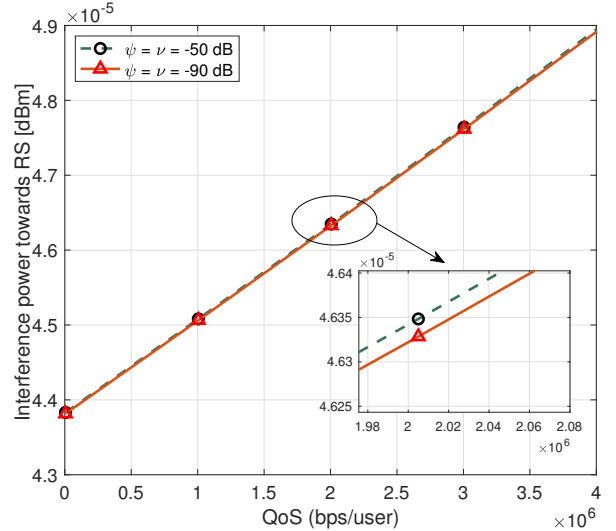


Fig. 5: Interference power towards RS vs. QoS per user at CS.

users and RS. To model the SI channel, the Rician model in [20] is adopted, wherein the SI channel is distributed as $\tilde{\mathbf{H}}_0 \sim \mathcal{CN}\left(\sqrt{\frac{K_R}{1+K_R}}\hat{\mathbf{H}}_0, \frac{1}{1+K_R}\mathbf{I}_{N_0} \otimes \mathbf{I}_{M_0}\right)$, where K_R is the Rician factor and $\hat{\mathbf{H}}_0$ is a deterministic matrix¹². Unless otherwise stated, we consider the following parameters for the CS and RS. For CS: thermal noise density = -174 dBm/Hz, noise figure at BS (users) 13(9) dB, $\tilde{N} = 2$, $\psi = \nu = -70$ dB, $\delta = \theta = 0.1$, $R_{i,min}^{UL} = R_{j,min}^{DL} = 5.0 \times 10^5$ bps, $P_i = 5$ dB, $P_0 = 10$ dB and CCI cancellation factor¹³ = 0.5. For RS:

¹²For simplicity, we take $K_R = 1$ and the matrix $\tilde{\mathbf{H}}_0$ of all ones for all simulations [22].

¹³It is essential to isolate UL and DL users in a FD system through smart channel assignments at a stage prior to the precoder/decoder design, so that the CCI is mitigated. This can be done by clustering the users into different groups through techniques, such as game theory, where the users with very strong CCI are not placed in the same group. In this work, the value of CCI cancellation factor represents the following: 0 \rightarrow 100% cancellation and 1 \rightarrow 0% cancellation.

$R = 8$, $P_{FA} = 10^{-5}$, velocity of target = 782 knots, and distance of target from RS = 300 m.

2) *Simulation Results:* To enable spectrum sharing, RS deploys Algorithm 1 to null the interference from RS \rightarrow CS, i.e., $\mathbf{W}_i \mathbf{x}_R = \mathbf{0}$, $i \in \mathcal{S}$, which paves the way for the CS to use the RS's spectrum. This is simultaneously followed by the deployment of Algorithm 2 at CS, which maximises the PoD of RS by suppressing the interference from CS \rightarrow RS, while also providing QoS to its users. In the following examples, we illustrate the performance of both RS and CS under a spectrum sharing scenario utilising the proposed algorithms. Due to the iterative nature of Algorithm 2, we begin by showing 1) its evolution in Fig. 2, i.e., its convergence and 2) its complexity analysis in Fig. 3 in terms of complex multiplications required with respect to (w.r.t) increasing number of antennas at CS and RS and users in the CS. It can be seen from Fig. 2 that the cost function, i.e., I^{RAD} monotonically decreases and converges after 25 – 30 iterations. Further, in Fig. 3, the axes in red (left

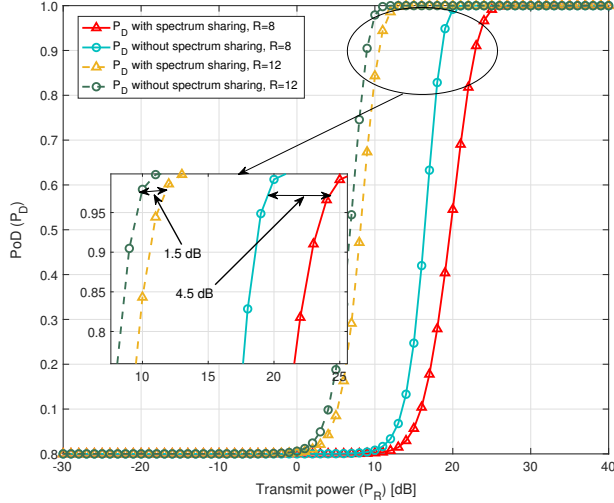


Fig. 6: PoD of MIMO RS vs. MIMO RS's transmit power.

and bottom) represent the complexity w.r.t. number of users in the CS, while the axes in black (right and top) represent the complexity w.r.t. number of antennas at RS and the FD BS. It can be seen that the computational complexity of Algorithm 2 increases as the number of users or antennas are increased. Hence, it is imperative that the processing of Algorithm 2 be handled centrally at the FD BS, which inadvertently has high end computing capabilities.

Next, we analyse the impact of spectrum sharing on the proposed two-tier model. In particular, we quantify the level of interference towards the RS for various levels of QoS that the CS can support by operating in FD mode. Accordingly, in Fig. 4 we show the interference power generated from the CS towards the RS as a function of transmitter/receiver (ψ/ν) distortion values for two different QoS requirements of the cellular users. Note that the transmitter/receiver distortion values reflect the amount of RSI left in the FD system. It can be seen from the figure that, as the RSI cancellation capability of the FD system increases, the interference power generated by the CS towards the RS decreases. This can be explained due to the fact that, when RSI is more, the CS has to use higher transmission power to overwhelm the RSI and maintain the QoS of the users, which results in more interference towards the RS. More importantly, it can also be seen that the minimum guaranteed QoS for each user can be increased by a factor of $10\times$ or more if the RS increases its tolerance threshold for interference temperature by only 0.6 dB. Similarly, in Fig. 5 we show the effect of users' QoS for two different RSI values. It can be seen that as the QoS requirements of cellular users increases, the interference from CS towards RS increases linearly. This can be explained due to the fact that, to provide higher QoS to the users, Algorithm 2 ensures transmission at higher power at the CS. Nevertheless, Algorithm 2 also ensures that for any specific QoS on the x -axis of Fig. 5, the corresponding y -axis value represents the minimum interference that can be generated from CS towards the RS.

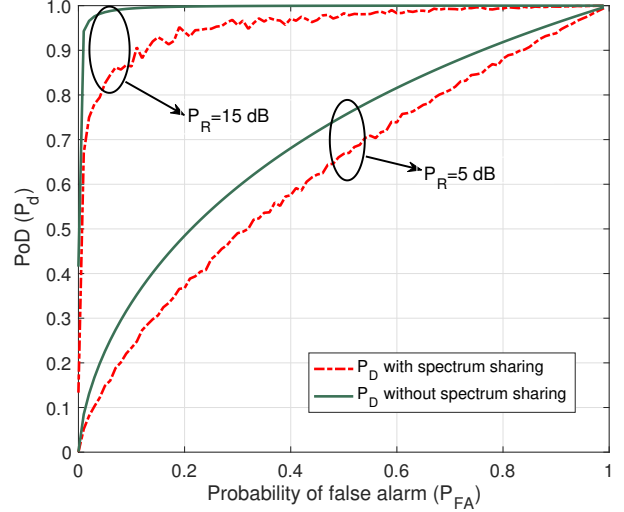


Fig. 7: PoD of MIMO RS vs. probability of false alarm at MIMO RS.

After quantifying the interference towards RS, we now determine its PoD. To detect a target in the far-field, the RS transmits NSP waveforms generated in Algorithm 1 and estimates parameters ϕ and α_r from the received signal that also includes I^{RAD} (obtained from Algorithm 2). As a benchmark, we also simulate the scenario without spectrum sharing by generating orthogonal waveforms at the RS and setting $I^{RAD} = 0$. This scenario relates to the case when the CS's BS is unable to provide its users with any connectivity due to lack of spectrum resources.

Accordingly, in Fig. 6, the PoD of the MIMO RS w.r.t. RS's transmit power is shown. Here, we consider two scenarios: 1) $R = 8$ (straight lines) and 2) $R = 12$ (dashed lines). It can be seen that for fixed P_{FA} , in order to achieve a particular P_D the RS needs more power (to create NSP waveforms and withstand interference from CS to enable spectrum coexistence) than the case without spectrum sharing scenario. However, it can be seen that the RS needs more power when $R = 8$ than $R = 12$ to achieve similar performance. This is because, while the number of antennas at the CS (BS and users) are fixed, increasing the RS's antennas means that it has more degrees of freedom for reliable target detection and simultaneously nulling out interference towards the CS. This proves that large antenna arrays can be used at the RS to facilitate spectrum sharing without any significant degradation in RS's performance.

Finally, in Fig. 7, we plot PoD for various P_{FA} and P_R . Similar to the previous figure, the PoD of the RS is better at high P_R and small P_{FA} when the RS is not sharing its spectrum. However when both P_R and P_{FA} are small, PoD of RS for NSP waveforms is quite comparable to the case without spectrum sharing.

VIII. CONCLUSION

A two tier spectrum sharing framework was proposed, where 1) transceivers were jointly designed at a hardware impaired FD CS under imperfect CSI considerations 2) null-

space based waveforms were designed at MIMO RS under perfect CSI considerations. In particular, the robust optimisation in the CS led to an intractable problem, which was transformed into an equivalent tractable semidefinite programming problem. Next, algorithms were proposed to suppress interference at both systems, thus maximising the PoD of RS and also maintaining a specific QoS for each user in CS. Finally, numerical results were used to demonstrate the effectiveness of the proposed algorithms, albeit certain trade-offs in RS's transmit power, PoD, and QoS of the users. In particular, it was seen that to facilitate spectrum sharing, thereby providing the users of CS with QoS of 5×10^{-5} bps/user, the MIMO RS needs to spend an extra power of 1.5 – 4.5 dB depending on the number of antennas it uses. Overall, the designed framework provides the essential understanding for successful development of future cellular systems in conjunction with federal incumbents that can operate under same spectrum resources.

APPENDIX A PROOF OF PROPOSITION 1

In order to prove Proposition 1, we first need to show that \mathbf{P} is a projector. From (15), we have

$$\begin{aligned} \mathbf{P}^H &= (\mathbf{X}\tilde{\Omega}\mathbf{X}^H)^H = \mathbf{X}\tilde{\Omega}^H\mathbf{X}^H = \mathbf{P}, \\ \text{and } \mathbf{P}^2 &= \mathbf{X}\tilde{\Omega}\mathbf{X}^H\mathbf{X}\tilde{\Omega}\mathbf{X}^H = \mathbf{P}. \end{aligned} \quad (\text{A.1})$$

The above equation holds due to the fact that $\mathbf{X}\mathbf{X}^H = \mathbf{I}$ as they are orthogonal matrices and $\tilde{\Omega}^2 = \tilde{\Omega}$ by construction. Now in order to show that \mathbf{P} is a projector, we show $\mathbf{P}\mathbf{x} = \mathbf{x}$, if $\mathbf{x} \in \text{range}(\mathbf{P})$. In other words, for some \mathbf{w} , $\mathbf{x} = \mathbf{P}\mathbf{w}$. Hence, from the above and (A.1), we have

$$\mathbf{P}\mathbf{x} = \mathbf{P}(\mathbf{P}\mathbf{w}) = \mathbf{P}^2\mathbf{w} = \mathbf{P}\mathbf{w} = \mathbf{x}, \quad (\text{A.2})$$

$$\mathbf{P}(\mathbf{P}\mathbf{x} - \mathbf{x}) = \mathbf{P}^2\mathbf{x} - \mathbf{P}\mathbf{x} = \mathbf{0}. \quad (\text{A.3})$$

Hence, $\mathbf{P}\mathbf{x} - \mathbf{x} \in \text{null}(\mathbf{P})$, which shows that \mathbf{P} is a null-space projection matrix. Accordingly,

$$\mathbf{W}\mathbf{P}^H = \mathbf{R}\tilde{\Omega}\mathbf{X}^H\mathbf{X}\tilde{\Omega}\mathbf{X}^H = \mathbf{0}. \quad (\text{A.4})$$

which proves that \mathbf{P} is an orthogonal projection matrix onto the null-space of \mathbf{W} .

APPENDIX B PROOF OF LEMMA 2

Since $\mathbf{A}(\phi)\mathbf{P}\mathbf{P}^H\mathbf{A}^H(\phi)$ and $\hat{\chi}$ are positive-definite, we have

$$\begin{aligned} &\text{tr}(\mathbf{A}(\phi)\mathbf{P}\mathbf{P}^H\mathbf{A}^H(\phi)\hat{\chi}^{-1}\hat{\chi}), \\ &\leq \text{tr}(\mathbf{A}(\phi)\mathbf{P}\mathbf{P}^H\mathbf{A}^H(\phi)\hat{\chi}^{-1}) \text{tr}(\hat{\chi}). \end{aligned} \quad (\text{B.1})$$

Now, a lower bound for $\text{tr}(\mathbf{A}(\phi)\mathbf{P}\mathbf{P}^H\mathbf{A}^H(\phi)\hat{\chi}^{-1})$ follows as

$$\text{tr}(\mathbf{A}(\phi)\mathbf{P}\mathbf{P}^H\mathbf{A}^H(\phi)\hat{\chi}^{-1}) \geq \frac{\text{tr}(\mathbf{A}(\phi)\mathbf{P}\mathbf{P}^H\mathbf{A}^H(\phi))}{\text{tr}(\hat{\chi})}. \quad (\text{B.2})$$

Using the property $\text{tr}(\mathbf{A}_1 + \mathbf{B}_1) = \text{tr}(\mathbf{A}_1) + \text{tr}(\mathbf{B}_1)$, where \mathbf{A}_1 and \mathbf{B}_1 are square matrices and $\varphi = \text{tr}(\mathbf{P}\mathbf{P}^H)$, we can rewrite (B.2) under $R_T = R_R = R$ to obtain the desired result.

APPENDIX C PROOF OF LEMMA 4

To prove this lemma, we first construct $\text{tr}(\mathbf{B}_i^H\mathbf{E}_i\mathbf{B}_i)$ using (37) as

$$\begin{aligned} \text{tr}(\mathbf{B}_i^H\mathbf{E}_i\mathbf{B}_i) &= \|\mathbf{B}_i^H(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i})\|_F^2 \\ &+ \sum_{j \in \mathcal{S}, j \neq i} \|\mathbf{B}_i^H\mathbf{U}_i^H\mathbf{H}_{ij}\mathbf{V}_j\|_F^2 \\ &+ \sum_{j \in \mathcal{S}} \sum_{\ell \in \mathcal{D}_j^{(T)}} \psi \|\mathbf{B}_i^H\mathbf{U}_i^H\mathbf{H}_{ij}\mathbf{\Xi}_\ell\mathbf{V}_j\|_F^2 \\ &+ \sum_{j \in \mathcal{S}} \sum_{\ell \in \mathcal{D}_i^{(R)}} v \|\mathbf{B}_i^H\mathbf{U}_i^H\mathbf{\Xi}_\ell\mathbf{H}_{ij}\mathbf{V}_j\|_F^2 \\ &+ P_R \|\mathbf{B}_i^H\mathbf{U}_i^H\mathbf{W}_i\| + \sigma_i^2 \|\mathbf{B}_i^H\mathbf{U}_i^H\|_F^2, \end{aligned} \quad (\text{C.1})$$

where $\mathcal{D}_j^{(R)}$ and $\mathcal{D}_j^{(T)}$ denote the set $\{1 \cdots \tilde{N}_j\}$ and $\{1 \cdots \tilde{M}_j\}$, respectively, while $\mathbf{\Xi}_\ell$ represents a square matrix with zero elements, except for the ℓ -th diagonal element, which is equal to 1. Utilising the $\text{vec}(\cdot)$ operation, and the identity $\|\text{vec}(\mathbf{A})\|_2^2 = \text{tr}\{\mathbf{A}\mathbf{A}^H\}$, (C.1) from above can be reformulated as

$$\begin{aligned} \text{tr}(\mathbf{B}_i^H\mathbf{E}_i\mathbf{B}_i) &= \|\text{vec}(\mathbf{B}_i^H(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}))\|_2^2 \\ &+ \sum_{j \in \mathcal{S}, j \neq i} \|\text{vec}(\mathbf{B}_i^H\mathbf{U}_i^H\mathbf{H}_{ij}\mathbf{V}_j)\|_2^2 \\ &+ \sum_{j \in \mathcal{S}} \sum_{\ell \in \mathcal{D}_j^{(T)}} \psi \|\text{vec}(\mathbf{B}_i^H\mathbf{U}_i^H\mathbf{H}_{ij}\mathbf{\Xi}_\ell\mathbf{V}_j)\|_2^2 \\ &+ \sum_{j \in \mathcal{S}} \sum_{\ell \in \mathcal{D}_i^{(R)}} v \|\text{vec}(\mathbf{B}_i^H\mathbf{U}_i^H\mathbf{\Xi}_\ell\mathbf{H}_{ij}\mathbf{V}_j)\|_2^2 \\ &+ P_R \|\text{vec}(\mathbf{B}_i^H\mathbf{U}_i^H\mathbf{W}_i)\|_2^2 + \sigma_i^2 \|\text{vec}(\mathbf{B}_i^H\mathbf{U}_i^H)\|_2^2. \end{aligned} \quad (\text{C.2})$$

By utilising the identity $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$, (C.2) in the above can be rewritten as $\|\mathbf{z}_{ij}\|_2^2$, where \mathbf{z}_{ij} is defined as

$$\mathbf{z}_{ij} = \begin{bmatrix} \text{vec}(\mathbf{B}_i^H(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i})) \\ \lfloor (\mathbf{V}_j^T \otimes (\mathbf{B}_i^H\mathbf{U}_i^H)) \text{vec}(\mathbf{H}_{ij}) \rfloor_{j \in \mathcal{S}, j \neq i} \\ \lfloor \lfloor \sqrt{\psi} ((\mathbf{\Xi}_\ell\mathbf{V}_j)^T \otimes (\mathbf{B}_i^H\mathbf{U}_i^H)) \text{vec}(\mathbf{H}_{ij}) \rfloor_{\ell \in \mathcal{D}_j^{(T)}} \rfloor_{j \in \mathcal{S}} \\ \lfloor \lfloor \sqrt{v} (\mathbf{V}_j^T \otimes (\mathbf{B}_i^H\mathbf{U}_i^H\mathbf{\Xi}_\ell)) \text{vec}(\mathbf{H}_{ij}) \rfloor_{\ell \in \mathcal{D}_i^{(R)}} \rfloor_{j \in \mathcal{S}} \\ P_R (\mathbf{I}_R^T \otimes (\mathbf{B}_i^H\mathbf{U}_i^H)) \text{vec}(\mathbf{W}_i) \\ \sigma_i \text{vec}(\mathbf{B}_i^H\mathbf{U}_i^H) \end{bmatrix} \quad (\text{C.3})$$

In a similar way, we can also represent I^{RAD} into vector form as

$$\begin{aligned} I^{RAD} &= \sum_{j \in \mathcal{S}} (\|\text{vec}(\mathbf{G}_{lj}\mathbf{V}_j)\|_2^2 \\ &+ \sum_{\ell \in \mathcal{D}_j^{(T)}} \psi \|\text{vec}(\mathbf{G}_{lj}\mathbf{\Xi}_\ell\mathbf{V}_j)\|_2^2). \end{aligned} \quad (\text{C.4})$$

Now, applying the same identity $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$, (C.4) in above can be constructed as $\|\boldsymbol{\iota}\|_2^2$, where $\boldsymbol{\iota}$ is defined as

$$\boldsymbol{\iota} = \begin{bmatrix} [(\mathbf{V}_j^T \otimes \mathbf{I}_T) \text{vec}(\mathbf{G}_j)]_{j \in \mathcal{S}} \\ \sqrt{\psi} [((\boldsymbol{\Xi}_\ell \mathbf{V}_j)^T \otimes \mathbf{I}_T) \text{vec}(\mathbf{G}_j)]_{\ell \in \mathcal{D}_j^{(T)}} \bigg]_{j \in \mathcal{S}}. \quad (\text{C.5})$$

APPENDIX D USEFUL LEMMAS

Lemma 5: If \mathbf{P} , \mathbf{Q} , \mathbf{A} are matrices with $\mathbf{A} = \mathbf{A}^H$, then semi-infinite LMI of the form of $\mathbf{A} \succeq \mathbf{P}^H \mathbf{X} \mathbf{Q} + \mathbf{Q}^H \mathbf{X}^H \mathbf{P}$, $\forall \mathbf{X} : \|\mathbf{X}\|_F \leq \rho$, holds iff $\exists \epsilon \geq 0$ such that

$$\begin{bmatrix} \mathbf{A} - \epsilon \mathbf{Q}^H \mathbf{Q} & -\rho \mathbf{P}^H \\ -\rho \mathbf{P} & \epsilon \mathbf{I} \end{bmatrix} \succeq 0. \quad (\text{D.1})$$

Lemma 6: Schur Complement Lemma [31]: Let \mathbf{Q} and \mathbf{R} are symmetric matrices. Then

$$\begin{bmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^* & \mathbf{R} \end{bmatrix} \succeq 0 \triangleq \mathbf{R} \succeq 0, \mathbf{Q} - \mathbf{S} \mathbf{R}^{-1} \mathbf{S}^* \succeq 0. \quad (\text{D.2})$$

REFERENCES

- [1] Federal Communications Commission (FCC), *Spectrum policy task force*, [Online] Available: <https://transition.fcc.gov/sptf/files/IPWGFFinalReport.pdf>, Nov. 2002.
- [2] Federal Communications Commission (FCC), *FCC proposes innovative small cell use in 3.5 GHz band*, [Online] Available: <https://www.fcc.gov/document/fcc-proposes-innovative-small-cell-use-35-ghz-band>, accessed, Dec. 2012.
- [3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.
- [4] CEPT, "Licensed Shared Access (LSA)," *Electronic Communications Committees (ECC)*, [Online] Available: <http://www.eroadocdb.dk/Docs/doc98/official/pdf/ECCREP205.PDF>, Feb. 2014.
- [5] M. Matinmikko, M. Mustonen, D. Roberson, J. Paavola, M. Hoyhtya, S. Yrjola, and J. Roning, "Overview and comparison of recent spectrum sharing approaches in regulation and research: From opportunistic unlicensed access towards licensed shared access," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DYSpan)*, Apr. 2014, pp. 92-102.
- [6] National Telecommunications and Information Administration (NTIA), *An assessment of the near-term viability of accommodating wireless broadband systems in the 1675-1710 MHz, 1755-1780 MHz, 3500-3650 MHz, 4200-4220 MHz, and 4380-4400 MHz bands (Fast track report)*, [Online] Available: https://www.ntia.doc.gov/files/ntia/publications/fasttrackevaluation_111-52010.pdf, Nov. 2010.
- [7] M. J. Marcus, "Sharing government with private users: Opportunities and challenges," *IEEE Wireless Commun.* vol. 16, no. 3, pp. 4-5, Jun. 2009.
- [8] F. Hesar and S. Roy, "Spectrum sharing between a surveillance radar and secondary Wi-Fi networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 3, pp. 1434-1447, Jul. 2016.
- [9] R. Saruthirathanaworakun, J. M. Peha, and L. M. Correia, "Opportunistic sharing between rotating radar and cellular," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 10, pp. 1900-1910, Nov. 2012.
- [10] A. Babaei, W. H. Tranter, and T. Bose, "A nullspace-based precoder with subspace expansion for radar/communications coexistence," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2013, pp. 3487-3492.
- [11] F. Liu, C. Masouros, A. Li, T. Ratnarajah, and J. Zhou, "MIMO radar and cellular coexistence: A power-efficient approach enabled by interference exploitation," *IEEE Trans. Signal Process.*, vol. 66, no. 14, pp. 3681-3695, Jul. 2018.
- [12] B. Li and A. Petropulu, "MIMO radar and communication spectrum sharing with clutter mitigation," in *Proc. IEEE Radar Conf. (RadarConf)*, May 2016, pp. 1-6.
- [13] B. Li and A. P. Petropulu, "Joint transmit designs for coexistence of MIMO wireless communications and sparse sensing radars in clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 6, pp. 2846-2864, Dec. 2017.
- [14] A. Khawar, A. Abdelhadi, and T. C. Clancy, "Target detection performance of spectrum sharing MIMO radar," *IEEE Sensors J.*, vol. 15, no. 9, pp. 4928-4940, Sept. 2015.
- [15] B. Li, A. Petropulu, and W. Trappe, "Optimum co-design for spectrum sharing between matrix completion based MIMO radars and a MIMO communication system," *IEEE Trans. Signal Process.*, vol. 64, no. 17, pp. 4562-4575, Sept. 2016.
- [16] J. A. Mahal, A. Khawar, A. Abdelhadi, and T. C. Clancy, "Spectral coexistence of MIMO radar and MIMO cellular system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 2, pp. 655-668, Apr. 2017.
- [17] H. Deng and B. Himed, "Interference mitigation processing for spectrum-sharing between radar and wireless communication systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 3, pp. 1911-1919, Jul. 2013.
- [18] A. C. Cirik, S. Biswas, S. Vuppala and T. Ratnarajah, "Robust transceiver design for full duplex multiuser MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 5, no. 3, pp. 260-263, Jun. 2016.
- [19] A. C. Cirik, S. Biswas, S. Vuppala and T. Ratnarajah, "Beamforming design for full-duplex MIMO interference channels-QoS and energy-efficiency considerations," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4635-4651, Nov. 2016.
- [20] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296-4307, Dec. 2012.
- [21] W. Li, J. Lilleberg, and K. Rikkinen, "On rate region analysis of half- and full-duplex OFDM communication links," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1688-1698, Sept. 2014.
- [22] D. Nguyen, L. Tran, P. Pirinen, and M. Latva-aho, "On the spectral efficiency of full-duplex small cell wireless systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4896-4910, Sept. 2014.
- [23] A. C. Cirik, S. Biswas, S. O. Taghizadeh, and T. Ratnarajah, "Robust transceiver design in full-duplex MIMO cognitive radios," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1-1, 2017.
- [24] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3702-3713, Jul. 2012.
- [25] E. A. Gharavol, Y.-C. Liang, and K. Mouthaan, "Robust downlink beamforming in multiuser MISO cognitive radio networks with imperfect channel-state information," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2852-2860, Jul. 2010.
- [26] T. W. Ban, W. Choi, B. C. Jung, and D. K. Sung, "Multi-user diversity in a spectrum sharing system," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 102-106, Jan. 2009.
- [27] L. Xu and J. Li, "Iterative generalized-likelihood ratio test for MIMO radar," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2375-2385, Jun. 2007.
- [28] I. Bekkerman and J. Tabrikian, "Target detection and localization using MIMO radars and sonars," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3873-3883, Oct. 2006.
- [29] S. M. Kay, *Fundamentals of statistical signal processing: Detection theory*, vol. 2, Prentice Hall, 1998.
- [30] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1-182, Jun. 2004.
- [31] S. Boyd and L. Vandenberghe, *Convex optimisation*, Cambridge, U.K.: Cambridge University Press, 2004.
- [32] Y. Zhang, E. Dall'Anese, and G. B. Giannakis, "Distributed optimal beamformers for cognitive radios robust to channel uncertainties," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6495-6508, Dec. 2012.
- [33] J. Jose, N. Prasad, M. Khojastepour, and S. Rangarajan, "On robust weighted-sum rate maximization in MIMO interference networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2011, pp. 1-6.
- [34] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex optimisation: Analysis, Algorithms, Engineering Applications*, Philadelphia, PA, USA: SIAM, 2001.
- [35] H. Shen, B. Li, M. Tao, and X. Wang, "MSE-based transceiver designs for the MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3480-3489, Nov. 2010.
- [36] J. B. Andersen, T. S. Rappaport and S. Yoshida, "Propagation measurements and models for wireless communications channels," *IEEE Commun. Mag.*, vol. 33, no. 1, pp. 42-49, Jan. 1995.
- [37] S. Sun et al., "Propagation path loss models for 5G urban micro- and macro-cellular scenarios," in *Proc. IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016, pp. 1-6.